

Université de Montréal
Faculté des Études Supérieures

Théorie et pratique
de la construction humaine supervisée du sens

Par :
Khalid Rouane

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade Philosophiae Doctor (Ph.D.)
en informatique

Juillet, 2004
(c) Khalid Rouane, 2004



QA

76

U54

2004

v. 039

Direction des bibliothèques

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
la Faculté des études supérieures

Cette thèse intitulée :
Théorie et pratique
de la construction humaine supervisée du sens

présentée par:
Khalid Rouane

a été évaluée par un jury composé des personnes suivantes :

Jean Meunier, président-rapporteur
Claude Frasson, directeur de recherche
Marc Kaltenbach, co-directeur de recherche
Mostapha Aboulhamid, membre du jury
Stefano A. Cerri, examinateur externe
Gilbert Babin, représentant du doyen de la FES

RÉSUMÉ

La e-Formation a pour objectif l'utilisation des moyens électroniques pour dispenser une formation individualisée à une grande audience. Des facteurs multiples comme l'omniprésence de l'accès Web par Internet et des ordinateurs personnels, accompagné d'une demande en forte croissance pour la formation due aux passages de la société d'une société *industrielle* à une société du *savoir*, ont mis la e-Formation sur le devant de la scène pour devenir un enjeu national dans la plupart des pays.

Une discipline importante dans la e-Formation est les *Systèmes Tutoriels Intelligents*, dits STIs. Leur objectif est la création de systèmes de soutien à l'apprentissage capables d'une interactivité et adaptabilité améliorées par l'utilisation des techniques d'intelligence artificielle. Le problème central dans cette discipline, comme c'est le cas dans l'intelligence artificielle en général, est la modélisation de la connaissance.

La majorité des travaux de recherches dans les STIs ont focalisé sur la modélisation de la connaissance qui sert principalement les opérations d'inférence et de traitement internes qu'effectuent ces systèmes. Ceci a mené à plusieurs difficultés dont les plus importantes sont : 1) la limitation de l'utilisation des STIs actuels aux sciences dures et aux domaines formels comme par exemple l'algèbre et la géométrie, et 2) l'absence de support effectif à l'apprenant du fait que la connaissance modélisée n'a pas été pensée dans ce sens.

Cette thèse propose des solutions pratiques à ces difficultés par l'utilisation d'une nouvelle approche dans la modélisation de la connaissance dans les STIs. Dans cette approche, la connaissance modélisée sert en premier lieu l'apprenant (et non le système) et spécialement lors de la lecture didactique des documents et supports textuels. La lecture didactique est une phase très critique dans le processus d'apprentissage car elle est le premier siège des opérations cognitives d'acquisition et d'organisation de la connaissance qu'effectue un apprenant. Des modèles de représentation de la connaissance sont proposés pour prendre en

compte les aspects reliés au discours, à l'épistémique et à la didactique. La modélisation du discours permet de mettre en évidence l'organisation et la structuration du discours contenu dans le texte didactique dans le but de faciliter la compréhension. La modélisation épistémique permet de représenter des abstractions, des concepts et des idées exprimées dans le texte sous forme graphique pour donner à l'apprenant une perspective supplémentaire sur le sujet traité. La modélisation didactique permet de placer l'expérience de la lecture dans un cadre pédagogique planifié et organisé.

Grâce à ces modèles, nous avons proposé de nouvelles méthodes pour la modélisation de l'apprenant et notamment son modèle cognitif et nous avons conçu de nouveaux mécanismes pour la génération de cours, basés sur les notions d'objectifs de compréhension et de sujets génératifs à maîtriser.

La thèse contribue au développement de système STIs pratiques et fonctionnels basés sur la lecture didactique. Ces systèmes transforment la phase de la lecture conventionnelle, traditionnellement très passive, en une lecture constructiviste où l'apprenant est très impliqué et sollicité. Ceci permet de créer des cours de qualité rapidement et plus facilement à partir de la gigantesque base de documents didactiques textuels déjà disponibles.

Mots-clés :

- système tutoriel intelligent (STI),
- représentation et modélisation de connaissances dans les STIs,
- représentations externes et concept-mapping,
- annotations,
- formes épistémiques,
- apprentissage et lecture didactique,
- modèle de l'apprenant.

ABSTRACT

The e-Learning aims to the use of electronic means to provide an individualized training to a large population of learners. Multiple factors such as the ubiquitous access to the Web by the Internet and powerful micro-computers, along with a strong growing demand for training due to the transformation of our society from an industrial society to a society of knowledge, the e-Learning has become a national priority in almost all industrialized countries.

An important discipline in the e-Learning is the Intelligent Tutoring Systems (ITS). Their objective is the creation of training systems capable of a great interactivity and adaptability thanks to the use of artificial intelligence techniques. The central problem in this discipline, as it is the case in artificial intelligence in general, is the knowledge modeling issue.

In the majority of ITS systems, modeled knowledge is mainly used in the internal inferences of these systems and this situation has carried out to several difficulties of which the most important are: 1) the limitation of the use of the current ITS to hard sciences and formal domains of knowledge, such as algebra and geometry, and 2) the absence of an effective support to the learner because the modeled knowledge was not thought to be used by the learner.

This thesis proposes practical solutions to these difficulties by the use of a new approach in the modeling of knowledge. In this new approach, modeled knowledge is aimed to be primarily used by the learner (and not the system) and especially at the time of the didactic reading of documents and textual medias. The didactic reading is a very critical phase in the process of training because it is the first seat of the cognitive operations of acquisition and organization of knowledge carried out by the learner. Models of knowledge are proposed that take into account aspects connected to the discourse, the epistemic and the didactic. Discourse modeling makes it possible to highlight the organization and the structuring of the discourse contained in a didactic textual document, and it aims to facilitate the

comprehension. Epistemic modeling makes it possible to represent abstractions of concepts and ideas expressed in the text in a graphical form to give the learner an additional perspective on the covered subject. Didactic modeling makes it possible to put the reading experience within a planned and organized teaching framework.

Using these models, we propose new methods for the learner modeling and particularly for the learner cognitive model. And we designed new mechanisms for the automatic generation of courses based on the concepts of understanding objectives and generative topics.

The thesis contributes to the development of practical and functional ITS based on the didactic reading. This kind of ITS transforms the phase of the conventional reading, which is very passive, into a constructivist reading where the learner participates actively. This makes it possible to create more courses of quality, quickly and easily, by exploiting the gigantic base of didactic textual documents already available.

Keywords:

- Intelligent tutoring systems (ITS),
- Knowledge modeling and representation in ITS,
- External representation and concept-mapping,
- Annotations,
- Epistemic forms,
- Learning and didactic reading,
- Learner modeling.

TABLE DES MATIÈRES

| | | |
|------------|--|----|
| Chapitre 1 | Introduction..... | 1 |
| 1.1 | But de la thèse | 5 |
| 1.2 | Organisation de la thèse..... | 6 |
| Chapitre 2 | Systèmes tutoriels intelligents..... | 7 |
| 2.1 | Des systèmes d'EAO traditionnels aux systèmes STIs | 7 |
| 2.2 | Structure générale d'un STI..... | 12 |
| 2.2.1 | Le module expert | 13 |
| 2.2.2 | Module tuteur | 17 |
| 2.2.3 | Modèle de l'apprenant..... | 20 |
| 2.2.4 | Interface de communication avec l'apprenant..... | 25 |
| 2.3 | Difficultés et critiques des approches STI actuelles..... | 26 |
| 2.3.1 | Difficultés sur le plan de la théorie..... | 26 |
| 2.3.2 | Difficultés dans la modélisation du domaine..... | 28 |
| 2.3.3 | Difficultés dans la modélisation de l'apprenant..... | 30 |
| 2.3.4 | Difficultés dans la pédagogie..... | 31 |
| 2.4 | Conclusion | 33 |
| Chapitre 3 | La conception pédagogique et les mécanismes d'apprentissage | 34 |
| 3.1 | La conception pédagogique : du behaviorisme au constructivisme | 34 |
| 3.1.1 | L'approche behavioriste..... | 34 |
| 3.1.2 | L'approche cognitiviste | 36 |

| | | |
|------------|--|----|
| 3.1.3 | L'approche constructiviste..... | 39 |
| 3.1.4 | Béhaviorisme, cognitivisme et constructivisme..... | 41 |
| 3.2 | Conception pédagogique et théories d'apprentissage..... | 43 |
| 3.2.1 | La théorie des conditions d'apprentissage de Gagné..... | 44 |
| 3.2.2 | La théorie des transactions d'enseignement de Merrill | 48 |
| 3.3 | Les mécanismes d'apprentissage..... | 51 |
| 3.3.1 | Apprentissage et stratégies de traitement de l'information..... | 51 |
| 3.3.2 | Apprentissage et représentation de la connaissance | 52 |
| 3.3.2.1 | Les concepts et les structures conceptuelles..... | 53 |
| 3.3.2.2 | Les schémas..... | 54 |
| 3.3.2.3 | Les théories et les modèles mentaux..... | 56 |
| 3.4 | Connaissance du domaine et apprentissage de haut niveau. | 59 |
| 3.4.1 | Nature de la connaissance du domaine..... | 59 |
| 3.4.2 | Apprentissage de haut niveau..... | 62 |
| 3.5 | Conclusion | 65 |
| Chapitre 4 | Le modèle <i>LEKC</i> : pour le discours, l'épistémique et la didactique | 66 |
| 4.1 | Introduction..... | 66 |
| 4.2 | Le modèle du discours (<i>LEKC-MD</i>)..... | 68 |
| 4.2.1 | La dimension propositionnelle..... | 69 |
| 4.2.2 | Les micropropositions..... | 70 |
| 4.2.3 | Les macropropositions..... | 71 |
| 4.2.4 | Modélisation formelle du modèle du discours..... | 74 |
| 4.2.5 | Utilité du modèle du discours | 75 |
| 4.3 | Le modèle épistémique (<i>LEKC-ME</i>) | 76 |
| 4.3.1 | Les dimensions de la modélisation épistémique..... | 77 |

| | | |
|------------|---|-----|
| 4.3.2 | La dimension rhétorique..... | 79 |
| 4.3.2.1 | Exemple de modélisation rhétorique..... | 79 |
| 4.3.2.2 | Les relations de la dimension rhétorique..... | 82 |
| 4.3.2.3 | Différence avec la RST classique..... | 85 |
| 4.3.2.4 | Limites de la modélisation rhétorique..... | 87 |
| 4.3.3 | La dimension thématique..... | 87 |
| 4.3.3.1 | Les types thématiques..... | 88 |
| 4.3.3.2 | Les relations thématiques..... | 89 |
| 4.3.3.3 | Exemple d'analyse thématique..... | 89 |
| 4.3.4 | La dimension structurelle..... | 90 |
| 4.3.5 | Modélisation formelle du modèle épistémique..... | 93 |
| 4.4 | Le modèle didactique (<i>LEKC-MI</i>)..... | 94 |
| 4.4.1 | Les sujets génératifs..... | 96 |
| 4.4.2 | Les objectifs de compréhension..... | 97 |
| 4.4.3 | Les activités épistémiques..... | 99 |
| 4.5 | Modélisation théorique de <i>LEKC</i> | 101 |
| 4.6 | Conclusion..... | 101 |
| Chapitre 5 | <i>IT-EKC</i> : une architecture de STI basé sur l'Approche <i>LEKC</i> | 103 |
| 5.1 | La lecture constructiviste..... | 104 |
| 5.2 | Connaissance modélisée et son utilisation..... | 105 |
| 5.3 | Vue globale du système <i>IT-EKC</i> | 106 |
| 5.4 | La composante d'édition et de conception..... | 107 |
| 5.4.1 | La modélisation au niveau discours..... | 109 |
| 5.4.2 | La modélisation au niveau épistémique..... | 110 |
| 5.4.3 | La modélisation au niveau contrôle..... | 112 |

| | | |
|------------|--|-----|
| 5.4.4 | La modélisation au niveau didactique | 117 |
| 5.4.5 | Diagramme des activités de conception | 118 |
| 5.5 | La composante de lecture assistée | 121 |
| 5.5.1 | Le module de visualisation et d'annotation | 122 |
| 5.5.2 | Le module planificateur | 123 |
| 5.5.3 | Le module tuteur | 124 |
| 5.5.4 | Le modèle de l'apprenant | 125 |
| 5.5.5 | L'éditeur des formes épistémiques | 128 |
| 5.5.6 | Diagramme des activités de la lecture assistée | 128 |
| 5.6 | La composante de collaboration | 133 |
| 5.6.1 | Collaboration synchrone | 134 |
| 5.6.2 | Collaboration asynchrone | 137 |
| 5.6.3 | Diagramme des activités de collaboration | 138 |
| 5.7 | Conclusion | 139 |
| Chapitre 6 | Apprentissage et enseignement basés sur <i>IT-EKC</i> | 140 |
| 6.1 | Exploitation de <i>IT-EKC</i> dans la création de cours | 141 |
| 6.1.1 | La notion de cours dans <i>IT-EKC</i> | 141 |
| 6.1.2 | L'approche de création de cours dans <i>IT-EKC</i> | 142 |
| 6.1.3 | Création de la méta-connaissance | 143 |
| 6.1.3.1 | Définition des méta-types | 143 |
| 6.1.3.2 | Création des méta-tags | 145 |
| 6.1.3.3 | Création des modèles <i>LEKC</i> | 148 |
| 6.1.4 | La spécification des besoins et des objectifs de formation | 150 |
| 6.1.5 | L'approche Model-based | 151 |
| 6.1.6 | L'approche Document-based | 154 |

| | | |
|------------|---|-----|
| 6.1.7 | La spécification du public cible..... | 155 |
| 6.1.8 | Personnalisation des objectifs de formation..... | 156 |
| 6.1.9 | Détermination du contenu à lire en fonction des objectifs de compréhension..... | 158 |
| 6.1.10 | Stratégies d'exploration du contenu..... | 162 |
| 6.2 | Stratégies d'interaction avec le système <i>IT-EKC</i> | 163 |
| 6.2.1 | Utilisation directe..... | 163 |
| 6.2.2 | Utilisation contrôlée..... | 165 |
| 6.2.2.1 | Stratégie du Model-tracing simple..... | 167 |
| 6.2.2.2 | Stratégie du Model-tracing avancé..... | 169 |
| 6.2.2.3 | Stratégie Configuration-based..... | 173 |
| 6.2.2.4 | Application des trois stratégies..... | 175 |
| 6.2.3 | Utilisation libre..... | 176 |
| 6.3 | Conclusion du chapitre..... | 178 |
| Chapitre 7 | Discussion et conclusion..... | 180 |
| 7.1 | Inconvénient de la stratégie de remédiation..... | 181 |
| 7.2 | Vers une stratégie de prévention..... | 182 |
| 7.3 | Autres avantages de l'approche <i>LEKC</i> | 183 |
| 7.4 | les limites de notre approche..... | 185 |
| 7.5 | Travaux futurs..... | 185 |
| | Bibliographie..... | 187 |
| | Annexe A Modélisation théorique de <i>LEKC</i> | 196 |
| | Annexe B Résumé de la notation BNF..... | 201 |

LISTE DES FIGURES

| | |
|---|----|
| Figure 2-1 – Architecture classique d'un STI..... | 13 |
| Figure 2-2 – Exemple de règle dans MYCIN | 15 |
| Figure 2-3 – Exemple de graphe de cours [Rouane, 1999]..... | 18 |
| Figure 2-4 – Réactions du système dans les stratégies Model Tracing et Issue-based..... | 20 |
| Figure 2-5 – Quatre vues de relations entre les connaissances de l'apprenant et..... | 23 |
| Figure 2-6 – Cycle des opérations dans l'activité d'enseignement | 32 |
| Figure 3-1 – Analogie avec le modèle de l'atome..... | 42 |
| Figure 3-2 – Circuit des neuf événements pédagogiques de Gagné..... | 47 |
| Figure 3-3 – Structure de réseau causal basé sur le modèle de Merrill..... | 50 |
| Figure 3-4 – Exemple de réseau sémantique [Collins et Quillian, 1969] | 54 |
| Figure 3-5 – Rôle de l'apprentissage dans la théorie des schémas..... | 56 |
| Figure 3-6 – Apprentissage et modèles mentaux | 57 |
| Figure 3-7 – Modèle mental pour la terre [Vosniadou et Brewer, 1992] | 58 |
| Figure 3-8 – Pyramide de la connaissance du domaine..... | 61 |
| Figure 3-9 – Apprentissage de bas niveau vs. de haut niveau | 64 |
| Figure 4-1 – Modèle de connaissance dans l'approche <i>LEKC</i> | 68 |
| Figure 4-2 – Exemple de décomposition en micropropositions | 71 |
| Figure 4-3 – Exemple de construction de macroproposition | 73 |
| Figure 4-4 - Définition du modèle du discours en notation BNF..... | 74 |
| Figure 4-5 – Diagramme de classe UML pour le modèle du discours..... | 75 |
| Figure 4-6 – L'abstraction et généralisation dans le modèle du discours | 76 |
| Figure 4-7 – Couplage des différentes dimensions | 78 |
| Figure 4-8 – Illustration d'une relation rhétorique de justification entre deux formes épistémiques..... | 79 |
| Figure 4-9 – Exemple simple d'analyse rhétorique (1) | 81 |
| Figure 4-10 – Exemple simple d'analyse rhétorique (2) | 82 |

| | |
|---|-----|
| Figure 4-11 – Exemple d’une analyse rhétorique complète [Mann, 2003] | 86 |
| Figure 4-12 – Un segment d’un texte scientifique..... | 90 |
| Figure 4-13 – Exemple simple d’analyse thématique..... | 90 |
| Figure 4-14 – Définition d’un model épistémique en notation BNF..... | 93 |
| Figure 4-15 – Diagramme de classe UML pour le modèle épistémique..... | 94 |
| Figure 4-16 – Modèle Didactique..... | 95 |
| Figure 4-17 – Organisation des sujets génératifs..... | 97 |
| Figure 4-18 – Organisation des objectifs de compréhension..... | 99 |
| Figure 4-19 – Organisation des activités épistémiques..... | 101 |
| Figure 5-1 – Proportion d’utilisation de la connaissance modélisée..... | 105 |
| Figure 5-2 – Utilisation de la connaissance modélisée proposée..... | 106 |
| Figure 5-3 – Les grandes composantes du système <i>IT-EKC</i> | 107 |
| Figure 5-4 – Schéma détaillé de la composante d’édition et de conception..... | 108 |
| Figure 5-5 – Exemple de patron épistémique pour l’argumentation..... | 112 |
| Figure 5-6 – Exemple d’instanciation d’un patron épistémique..... | 112 |
| Figure 5-7 – Relation entre les différents poids..... | 116 |
| Figure 5-8 – Liens entre les niveaux de modélisation..... | 118 |
| Figure 5-9 – Diagramme d’activités de conception..... | 120 |
| Figure 5-10 – Schéma détaillé de la composante de lecture assistée..... | 122 |
| Figure 5-11 – Exemple d’une séquence de planification..... | 124 |
| Figure 5-12 – La performance épistémique..... | 126 |
| Figure 5-13 – Modèle de cognitif apprenant..... | 127 |
| Figure 5-14 - Diagramme des activités de la lecture assistée..... | 131 |
| Figure 5-15 - Diagramme des activités utilisation libre..... | 132 |
| Figure 5-16 – Composante de collaboration..... | 133 |
| Figure 5-17 – Collaboration synchrone entre apprenants..... | 134 |
| Figure 5-18 – Diffusion des interactions..... | 136 |
| Figure 5-19 – Collaboration asynchrone..... | 137 |
| Figure 5-20 – Diagramme des activités de collaboration synchrone..... | 138 |
| Figure 5-21 – Diagramme des activités de collaboration asynchrone..... | 139 |
| Figure 6-1 – Interface de l’atelier de construction de cours..... | 141 |
| Figure 6-2 – Les composantes d’un cours dans <i>IT-EKC</i> | 143 |
| Figure 6-3 – Exemple de définition de méta-types nœuds..... | 144 |

| | |
|---|-----|
| Figure 6-4 – Outils de définition des méta-types nœuds..... | 144 |
| Figure 6-5 – Exemple de définition de méta-types liens..... | 144 |
| Figure 6-6 – Outils de définition des méta-types liens..... | 145 |
| Figure 6-7 – Navigateur et éditeur des méta-tags..... | 146 |
| Figure 6-8 – Exemple de code méta-tags HTML..... | 146 |
| Figure 6-9 – Reconstruction Dynamique des Méta-tags..... | 147 |
| Figure 6-10 – Interface d'édition des modèles de domaine..... | 149 |
| Figure 6-11 – Interface d'édition du modèle didactique..... | 149 |
| Figure 6-12 – Spécification des besoins de formation..... | 151 |
| Figure 6-13 – Approches de spécification des objectifs de formation..... | 151 |
| Figure 6-14 – Éditeur de définition de cours..... | 152 |
| Figure 6-15 – Mode d'association large et ciblée..... | 153 |
| Figure 6-16 – Éditeur de public cible..... | 156 |
| Figure 6-17 – Projection de S sur le document didactique..... | 159 |
| Figure 6-18 – Le “Quoi” et le “Comment” dans un cours..... | 160 |
| Figure 6-19 – Stratégie de présentation de contenu..... | 163 |
| Figure 6-20 – Bloc note interactif..... | 164 |
| Figure 6-21 – Interface tuteur <i>IT-EKC</i> | 165 |
| Figure 6-22 – Pyramide de la connaissance du domaine..... | 166 |
| Figure 6-23 – Explication supportant / s'opposant à l'insertion de α | 171 |
| Figure 6-24 – Interaction apprenant / système dans un dialogue pseudo-socratique..... | 173 |
| Figure 6-25 – Utilisation simultanée des trois stratégies d'intervention..... | 175 |
| Figure 6-26 – Application des trois stratégies et cycles du cours..... | 176 |
| Figure 6-27 – Base de connaissance collective..... | 177 |
| Figure 6-28 – Rôle du système dans la supervision..... | 178 |
| Figure 7-1 – Circuit pédagogique classique et la stratégie de remédiation..... | 181 |
| Figure 7-2 – Nouveau circuit pédagogique avec la stratégie de prévention..... | 182 |

LISTE DES TABLEAUX

| | |
|--|----|
| Table 2-i – Comparaison entre les STIs et les traditionnels systèmes d'E.A.O [Seidel et Park, 1994] | 10 |
| Table 2-ii – STIs historiques et leurs domaines d'application [Kaplan et Rock, 1995] | 12 |
| Table 3-i – Histoire de la relation entre la conception pédagogique et la psychologie éducative [Wilson et Cole, 1996] | 36 |
| Table 3-ii – Comparaison des théories behaviorisme, cognitivisme et constructivisme | 43 |
| Table 3-iii – Matrice des tâches | 49 |
| Table 3-iv – Définition des tâches épistémiques selon Ohlsson..... | 64 |

REMERCIEMENTS

Je tiens à remercier toutes les personnes qui ont contribué de près ou de loin à ce travail. En particulier, je tiens à remercier vivement mes directeurs de recherche Monsieur Claude Frasson (Université de Montréal) et Monsieur Marc Kaltenbach (Université Bishops). Leurs appuis tant intellectuels que financiers, sur toute la durée de ce projet qui a dépassé les cinq années, ont été la clé de son aboutissement. Leurs directives et leurs conseils, ainsi que leurs critiques, m'ont beaucoup inspiré et ont grandement influencé les orientations de mes travaux.

Je tiens à remercier chaleureusement tous les membres du jury et en particulier Monsieur Stefano A. Cerri pour tout le temps qu'il a consacré à la lecture et à l'évaluation de ma thèse. Ces remarques extrêmement pertinentes m'ont apporté un nouveau regard sur ma thèse et vont certainement influencer la suite de mes travaux. Par la même occasion je tiens à remercier Madame Esma Aïmeur pour toutes les discussions et échanges d'idées qu'on a entretenus durant ce projet.

Je tiens aussi à remercier mon épouse Souad pour sa patience, son support et sa compréhension. Elle a assumé la lourde tâche de prendre soin de nos deux enfants Sara et Ismail et de moi-même durant toute cette longue période. Par sa grande contribution indirecte à mon travail, cette thèse est aussi la sienne.

Je tiens finalement à remercier mon père Monsieur Rouane Mohammed et ma mère Madame Sidik Daouia, pour leur soutien moral durant toutes mes études.

À Sonad, Sara et Ismail.

CHAPITRE 1

INTRODUCTION

Les systèmes tutoriels intelligents, dits STIs, sont des systèmes d'enseignement informatisés ayant pour objectif de permettre une formation individualisée et adaptée à chaque apprenant. Ils vont plus loin que les traditionnels systèmes d'enseignement assisté par ordinateur (E.A.O.) par leur utilisation des techniques d'Intelligence Artificielle (IA) et leur modélisation des différentes connaissances impliquées dans le processus d'enseignement et d'apprentissage, à savoir, l'expertise du domaine, l'expertise pédagogique et la modélisation de l'apprenant [Murray, 1998-a].

La recherche dans ce domaine a été très fertile durant les deux dernières décennies. Cet intérêt peut s'expliquer par plusieurs raisons [Sandberg et Andriessen, 1999]. Premièrement, ces systèmes constituaient une évolution logique des anciens systèmes d'E.A.O., qui ne sont capables d'aucune adaptation ou personnalisation de la formation, ayant des scénarios d'enseignement figés et ne possédant aucune capacité de raisonnement. Deuxièmement, ces systèmes sont à la croisée de plusieurs disciplines scientifiques réunies, comme: la représentation de connaissance, le raisonnement, l'apprentissage machine, la planification, mais aussi, les sciences cognitives, les sciences de l'éducation, les interfaces homme--machine et bien d'autres.

Dès ses débuts dans les années quatre-vingt, la recherche dans le domaine des STIs s'est fixée des objectifs très ambitieux :

- formation individualisée adaptée au niveau de l'apprenant,
- suivi rapproché et analyse des activités de l'apprenant,

- interventions pertinentes et aux moments opportuns,
- aide à l'apprenant par génération d'explications adéquates,
- automatisation de la construction et de la génération de cours,
- incorporation de stratégies pédagogiques multiples.

Pour y arriver, les chercheurs en STI, très influencés par *l'hypothèse de la représentation symbolique* (suivant les travaux de Newell, A. et Simon, H. en intelligence artificielle [Luger et Stubblefield, 1998]), se sont efforcés de représenter le plus fidèlement possible, les connaissances des différents acteurs présents dans un processus d'apprentissage, qui sont l'expert du domaine, l'expert pédagogique et l'apprenant. Ceci a donné naissance à un paradigme qui est devenu classique et dominant dans le domaine des STIs [Polson et Richardson, 1988], basé sur le quadruple *{Modèle de la matière, Modèle pédagogique, Modèle de l'apprenant, Environnement-Interface}*.

Si ce paradigme avait démontré sa validité dans des domaines de connaissance fermés (dits micro-mondes [Quéré, 1991]), avec connaissance généralement de type procédurale et ayant une connaissance structurée par nature, comme dans les STIs *CABRI Géomètre* [Laborde, 1995] et *REFRACT* [Reimann, 1991], il s'est avéré, par contre, très difficilement généralisable à des domaines de connaissance plus ouverts [Sandberg et Andriessen, 1997 et 1999], aux connaissances peu ou mal structurées. Ce qui explique en grande partie, l'échec de ce paradigme [Du Boulay, 2000] et la quasi-absence jusqu'à aujourd'hui d'applications commerciales ou fonctionnelles, basées totalement sur ce modèle. Déjà en 1988, William Johnson [Polson et Richardson, 1988], préconisait que l'épreuve du monde réel est le test suprême que les STIs doivent passer :

“If ITSs are to reach their promise, then the laboratory systems must operate and survive in the real world.”

Ce même auteur, dans un sous-chapitre du même livre [Polson et Richardson, 1988], intitulé *"ITS Pragmatics"*, écrivait:

“...Certainly is true that nothing is as practical as good theory; nevertheless, there are *pragmatic issues* beyond good theory. *Who are the users?* What are the *expectations?* How can intelligent tutors be effectively implemented? Suffice it to say that a person does not simply decide in

vacuum that an ITS is the most appropriate means of instruction for a given domain, certain *practical matters* must be considered.”

Ces considérations d'ordre pratique et pragmatique ont été traitées sur deux pages seulement, dans un livre¹ de 280 pages, considéré comme la référence majeure dans le domaine des STIs. Une proportion qui montre combien on a sous-estimé l'importance de ces considérations et combien le rôle et les attentes de l'apprenant, pourtant unique *consommateur* de ces systèmes, ont été négligés. Des facteurs qui ont certainement contribué à entraver l'épanouissement des STIs en dehors des laboratoires, en dépit d'une demande, tant économique que sociale, sans cesse croissante.

Devant ce constat d'échec, la recherche dans les STIs a commencé ces dernières années à explorer de nouvelles voies en considérant d'avantage les questions d'ordre pratique et pragmatique et en s'appuyant sur les récents progrès accomplis dans les sciences cognitives et l'apprentissage.

Le paradigme classique, basé sur le modèle du système expert, hyper-intelligent, qui connaît *tout* de la matière et de l'apprenant, est progressivement abandonné au profit d'une nouvelle vision [Murray, 1998-b] qui cherche un équilibre entre:

- la puissance d'inférence des STIs, en d'autres mots leur intelligence, et par conséquent la quantité et le détail de la connaissance qui y est représentée,
- et la convivialité et la facilité d'emploi, tant pour les concepteurs de cours que pour les apprenants qui utilisent ces systèmes.

Des aspects très pratiques dans le processus d'apprentissage, comme la collaboration [Tadié, 1998] [Self, 1999] et la visualisation [Du Boulay, 2000], ont commencé à susciter beaucoup d'intérêt. Mais dans une optique plus pragmatique, beaucoup de points restent à considérer, surtout si on se met dans la position de l'apprenant qui va utiliser ces systèmes. Quels sont ses vrais besoins? Quelle aide effective le système peut-il lui apporter? Comment cette aide peut elle s'insérer harmonieusement et efficacement dans son travail quotidien d'étudiant? Peut-on aider l'étudiant à extérioriser ses idées et son raisonnement? Peut-on l'aider à les organiser, à les améliorer?

¹ [Polson et Richardson, 1988]

Le processus de l'éducation en général est composé de deux activités qui se déroulent en parallèle:

- une activité d'*enseignement*, menée par le tuteur (homme ou machine) via des outils ou média comme la parole, le texte, le schéma, etc.
- et une activité d'*apprentissage* menée par un apprenant avec, ce qu'on oublie souvent, des outils, comme la prise de notes, la construction de schémas, etc.

Les systèmes proposés par la recherche en STIs sont trop orientés (pour ne pas dire exclusivement orientés) sur des aspects de modélisation, de conception (*Authoring Tools*) et d'inférence. Le support à l'apprenant est presque inexistant dans ces systèmes. L'ajout d'un simple bloc-notes à un STI ne peut pas être considéré comme un support à l'apprenant, tant qu'il est considéré comme un élément très périphérique aux autres modules, presque décoratif et totalement opaque pour le système¹.

Selon les théories d'apprentissage les plus influentes ces dernières années, comme la théorie du constructivisme [Bruner, 1973], l'apprentissage est principalement une construction active de la connaissance par l'apprenant et non un simple transfert passif de cette connaissance, de l'expert à l'apprenant. La recherche dans le domaine des STIs s'est focalisée d'avantage sur les outils de modélisation de la connaissance et sur les mécanismes d'inférences exécutés par le système, au détriment des outils de support effectif à l'apprenant dans son effort pour construire, organiser et assimiler la nouvelle connaissance à apprendre.

De plus, les STIs actuels sont trop orientés vers la gestion de performance et sa mesure par les différentes activités de contrôles, telles la réponse aux questionnaires et la résolution de problèmes. Les erreurs détectées sont alors analysées et le système réagit pour remédier à la situation. C'est la stratégie de *l'enseignement correctif*, ('*Remediation Technique*', [Khuwaja, 1994]). Or les difficultés rencontrées par l'apprenant à ce niveau de contrôle, ont souvent leurs origines dans les erreurs de compréhension commises lors des phases initiales de construction mentale de la nouvelle connaissance par l'apprenant, comme durant la lecture des ressources pédagogiques. Ce qui démontre la nécessité et l'importance des outils de support à l'apprenant dans les phases de construction et d'organisation de la connaissance.

¹ Une boîte noire

1.1 But de la thèse

Notre principal objectif est de mettre au point une nouvelle approche pour la construction de systèmes STIs basés sur une *stratégie préventive* des erreurs de compréhension, spécialement dans les activités de lectures didactiques. Cette approche permet la conception et la réalisation d'outils et de méthodes pour assister efficacement l'apprenant dans les processus cognitifs d'acquisition et d'organisation de la connaissance lors de la lecture. Ces outils permettront à l'apprenant entre autres, de rendre explicite sa construction du sens et d'argumenter sur son raisonnement. L'introduction d'une telle approche, va nécessiter une révision complète des techniques de modélisation de la connaissance et des stratégies pédagogiques utilisées.

Nous nous proposons dans cette thèse:

- d'explorer les techniques de modélisation de la connaissance qui peuvent tenir compte de la spécificité du contexte de la lecture. Elles doivent prendre en compte des notions comme la structuration du discours, les relations rhétoriques, la thématique, ainsi que l'organisation didactique de la lecture ;
- d'explorer les théories d'apprentissage en se focalisant sur les notions de la compréhension et de l'apprentissage de haut niveau (ou de l'apprentissage complexe) et de réaliser des adaptations pour le contexte de la lecture didactique ;
- de trouver de nouvelles façons pour modéliser le profil cognitif de l'apprenant en fonction de la nouvelle modélisation de la connaissance proposée ;
- d'étudier les problèmes de l'évaluation et de l'analyse des représentations d'idées (les annotations) faites par l'apprenant lors de la lecture ;
- d'étudier la génération des dialogues par le système ainsi que les stratégies et les méthodes d'intervention qu'il peut déployer un tel contexte ;
- d'étudier les différentes façons de rendre la connaissance modélisée accessible à l'apprenant ;
- d'étudier les problèmes liés à la collaboration synchrone et asynchrone entre apprenants ;

- d'étudier le problème de la génération de cours dans un apprentissage basé sur la lecture didactique.

1.2 Organisation de la thèse

Dans le prochain chapitre nous allons exposer les difficultés liées aux fondements théoriques des STIs actuels ainsi que les problèmes rencontrés dans la modélisation du domaine et du modèle de l'apprenant. Nous examinerons aussi les difficultés relatives aux aspects pédagogiques.

Au troisième chapitre nous allons introduire des aspects importants de la conception pédagogique (*Instructional design*) et leur relations avec les théories et les mécanismes d'apprentissage. De même nous aborderons l'aspect de la représentation de la connaissance, mais sous un angle cognitif et non informatique.

Cet aspect de la représentation de la connaissance sera approfondi et plus en détail dans le chapitre quatre où nous allons présenter notre modèle *LEKC*¹. Ce modèle a été conçu spécialement pour supporter un apprentissage basé sur la lecture. Il prend en compte les aspects cognitifs de la compréhension du texte allant de la représentation du discours, en passant par la représentation schématique des raisonnements et des modèles mentaux, et en terminant par l'organisation didactique.

Au cinquième chapitre nous présenterons une architecture complète d'un système STI basé sur le modèle *LEKC*. Cette nouvelle architecture proposée, nommée *IT-EKC*², a été pensée pour supporter une lecture constructiviste dans le cadre d'une formation basée sur la lecture didactique. Nous verrons en détail les composantes d'édition et de conception, de lecture assistée et de collaboration. Nous présenterons au chapitre six une implantation de cette architecture, et nous décrirons les processus de création d'un cours dans cette approche ainsi que les stratégies et les méthodes d'interaction avec l'apprenant.

Nous concluons cette thèse au chapitre sept par une discussion sur les contributions et les limites de notre approche et par un aperçu sur les travaux futurs que nous comptons faire.

¹ pour "Learning by Explicit Knowledge Construction"

² pour "Intelligence Tutoring System based on Explicit Knowledge Construction"

CHAPITRE 2

SYSTÈMES TUTORIELS INTELLIGENTS

Les systèmes tutoriels intelligents, dits STIs, sont des applications didactiques informatisées qui visent une formation individualisée et adaptée à chaque apprenant. Ils sont conçus avec incorporation de techniques d'intelligence artificielle pour aboutir à des systèmes *tuteurs* qui savent *quoi* enseigner (l'expertise du domaine), *comment* l'enseigner (l'expertise pédagogique) et qui ont une connaissance de qui reçoit cet enseignement (la modélisation de l'apprenant) [Polson et Richardson, 1988] [Nwana, 1990].

L'objectif ultime de la réalisation de ces systèmes est le développement d'environnements d'enseignement hautement personnalisables dans lesquels l'apprenant et l'ordinateur peuvent avoir une flexibilité et interactivité très semblables à ce qui se passe quand un apprenant et un tuteur humain interagissent dans un processus conventionnel enseignement-apprentissage [Seidel et Park, 1994].

2.1 Des systèmes d'EAO traditionnels aux systèmes STIs

Les débuts de l'EAO

L'utilisation de l'ordinateur dans l'enseignement n'est pas propre aux STIs. L'enseignement assisté par ordinateur, communément appelé EAO, est une discipline scientifique qui a tenté d'utiliser l'ordinateur à des fins éducatives dès les débuts de l'ère informatique. Elle a

commencé à voir le jour il y a à peu près une cinquantaine d'années. Dans les années 50, Skinner [Skinner, 1958] a proposé un programme simple d'enseignement nommé "*Linear Programs*". Le matériel didactique dans ce programme était présenté pas-à-pas à l'apprenant dans une série de frames. La plupart des frames avaient une ou plusieurs questions simples avec des feedbacks immédiats du système et le programme passait d'une frame à l'autre sans prendre en considération la compréhension de l'apprenant dans la frame précédente. Ce type de programmes didactiques n'offre aucune individualisation de la formation; tous les apprenants utilisant un tel système recevaient exactement la même séquence de frames, sans tenir compte de leurs habiletés, background, et connaissances antérieures du domaine. [Carbonell, 1970] a critiqué ce type de programme en commentant que dans ce contexte, l'ordinateur ne fait qu'un peu mieux comparativement à un manuel programmé.

[Crowder, 1959] a proposé l'utilisation des réponses de l'apprenant pour contrôler la séquence d'affichage du matériel didactique. Le mécanisme de branchement ainsi introduit avait toujours un nombre fixe de frames possibles, mais il pouvait commenter les réponses de l'apprenant et les utilisait pour décider du frame suivante à afficher. La particularité de ce genre de programmes est qu'ils étaient capables d'offrir un feedback correctif et d'adapter la sélection du matériel présenté à l'apprenant.

Au début, les réponses étaient traitées en mode juste-ou-faux. Par la suite, des techniques de *pattern matching* ont permis que des réponses alternatives de l'apprenant soient traitées comme acceptables ou partialement acceptables, plutôt que totalement correctes ou erronées. Cependant, le matériel didactique devenait trop large pour être géré par une programmation directe. Ce qui a donné naissance à des langages de programmation spécialisés, appelés "*Authoring Languages*" [Nwana, 1990].

Les systèmes génératifs

Vers la fin des années 60 et le début des années 70, les *systèmes génératifs*, parfois appelés *systèmes adaptatifs*, ont commencé à voir le jour. L'idée est que des problèmes bien définis sont modélisés dans le système didactique en des instances multiples du même type de problèmes. L'apprenant, pour s'exercer à résoudre un type de problèmes donné, peut demander, la génération de nouvelles instances du même problème autant qu'il le désire. Le système calcule la solution et compare le travail de l'apprenant à cette solution. [Uhr, 1969] a

développé une série de systèmes didactiques qui généraient des problèmes en arithmétique avec des niveaux de difficultés variables et qui changeaient selon les performances de l'apprenant. Ce genre de systèmes était limité aux seuls exercices de pratique. Ils ne possédaient aucune connaissance réelle du domaine et ils ne pouvaient pas répondre aux questions que peut se poser l'apprenant. La différence était considérable entre les mécanismes internes de ces programmes et les processus cognitifs de l'apprenant dans la résolution des problèmes posés. Seules les interactions de l'apprenant étaient utilisées pour guider la génération de nouveaux problèmes, et aucune connaissance explicite sur l'apprenant n'y était représentée. Le modèle de l'apprenant était très rudimentaire, et il consistait parfois en un seul entier qui indiquait le niveau de l'apprenant.

Les débuts des STIs

[Self, 1974] suggérait qu'un système tutoriel devrait avoir une représentation de ce qui est enseigné, de comment l'enseigner et une représentation de qui reçoit cet enseignement (le modèle de l'apprenant). Le système devrait dynamiquement analyser l'historique des solutions et interactions avec l'apprenant, et utiliser des stratégies et des principes pour décider de la suite des actions à prendre, plutôt que de demander au concepteur d'anticiper et fournir à l'avance l'acheminement des actions que le système devait prendre. Cette vision pour les systèmes tutoriels a évolué avec l'exploitation des avancées dans le domaine de l'intelligence artificielle au milieu des années 80. Les systèmes d'EAO traditionnels sont devenus des EIAO pour *Enseignement Intelligemment Assisté par Ordinateur*. Ils ont reçu par la suite le nom distinctif de STI pour *Systèmes Tutoriels Intelligents* [Polson et Richardson, 1988]. Une comparaison des STIs et des systèmes EAO est donnée dans la Table 2-i.

| Comparaison | E.A.O. | ITSs |
|--------------------------|---|---|
| Objectifs de réalisation | Orienté vers des utilisations pratiques | Orienté vers l'exploration de nouvelles disciplines, surtout l'IA, dans l'éducation |
| Base théorique | Théories traditionnelles de l'apprentissage et de l'éducation | Théories de traitement de l'information et des sciences cognitives |

| Comparaison | E.A.O. | ITSs |
|--|---|---|
| Structures des systèmes et processus | Orienté frames statiques; structure unique; réactions préprogrammées du système | Orienté processus dynamiques; structure modulaire; réactions du système génératives |
| Principes pédagogiques | Variés, généralement à approche expositoire à l'initiative du système | Approche de découverte à l'initiative de l'apprenant |
| Méthodes de structuration de la connaissance | L'analyse de tâches principalement, pour l'identification de sous-tâches et des éléments de contenu | Techniques de représentation des connaissances |
| Méthode de modélisation de l'apprenant | Jugement et évaluation quantitative | Raisonnement et évaluation qualitative |
| Formats de l'instruction | Tutoriels généralement expositives avec des exercices | Tutoriels analytiques |
| Domaines d'application | Virtuellement tous les domaines | Limités aux domaines structurés et formels |
| Processus de développement et équipes | Approche système; concepteurs de cours, experts de la matière et programmeurs | Approche prototype; expert IA et ingénieur de la connaissance |

Table 2-i – Comparaison entre les STIs et les traditionnels systèmes d'E.A.O [Seidel et Park, 1994]

Il est difficile de donner une définition très précise de ce qu'est un STI, mais on peut approximer cette définition par la liste des caractéristiques suivantes :

- Un STI possède une connaissance assez détaillée du domaine à enseigner.
- Un STI possède une connaissance de l'apprenant de façon individuelle. Cette connaissance, nommée *modèle de l'apprenant*, est mise à jour dynamiquement et elle est utilisée pour guider l'instruction.
- Le concepteur d'un STI définit la connaissance du domaine et éventuellement les règles d'inférence, mais il ne définit pas la séquence d'enseignement. Cette séquence est

déterminée dynamiquement par le système en fonction des interactions avec l'apprenant et de son modèle.

- Un STI doit aller au-delà du simple contrôle de validité des réponses données par l'apprenant et doit être capable d'analyser et de diagnostiquer les causes des erreurs commises.
- Un STI doit être capable de répondre à des questions que peut se poser un apprenant. Ces questions peuvent être reliées soit au domaine de connaissance enseigné soit aux choix pédagogiques faits par le système.

En suivant plus au moins ces lignes directives dans la construction des STIs, plusieurs systèmes ont été construits et la Table 2-ii donne une liste des STIs les plus importants qui ont marqué l'histoire de cette discipline [Kaplan et Rock, 1995].

| ITS | Créateur(s) | Domaine |
|----------------|--------------------|-------------------------------|
| SCHOLAR | Carbonell | Géographie |
| WHY | Stevens, Collins | Météorologie |
| SOPHIE | Brown, Burton | Électronique |
| WUSOR | Goldstein | Jeux de stratégies |
| GUIDON | Clancey | Diagnostic médical avec Mycin |
| WEST | Burton | Jeux de stratégies |
| BUGGY | Brown | Arithmétique |
| DEBUGGY | Burton, VanLehn | Arithmétique |
| STEAMER | Stevens, Hollan | Moteurs de bateaux |
| LMS | Sleeman | Algèbre |
| MENO | Woolf | Météorologie |
| PROUST | Johnson | Programmation |
| ACTP | Anderson | Tuteur Lisp |

| ITS | Créateur(s) | Domaine |
|----------|---------------|--------------|
| SIERRA | VanLehn | Arithmétique |
| SHERLOCK | Lesgold, Katz | Électronique |

Table 2-ii – STIs historiques et leurs domaines d'application [Kaplan et Rock, 1995]

2.2 Structure générale d'un STI

Dans un processus d'enseignement-apprentissage réel, un enseignant humain utilise trois types de connaissances pour mener à bien son activité d'enseignement :

- une connaissance de la matière à enseigner,
- une connaissance des méthodes pédagogiques et
- une connaissance de l'apprenant qui reçoit cet enseignement.

Sur le même principe, Burns et Capps [Burns et Capps, 1988], ont posé les fondements de base pour une architecture d'un système STI (Figure 2-1), et qui depuis a rallié un grand consensus. Cette architecture stipule que tout STI doit avoir les quatre composantes suivantes:

- le module expert (ou l'expert du domaine)
- Le modèle de l'apprenant
- Le module tuteur
- L'interface apprenant

Les paragraphes qui suivront présenteront en détail chacune de ces composantes.

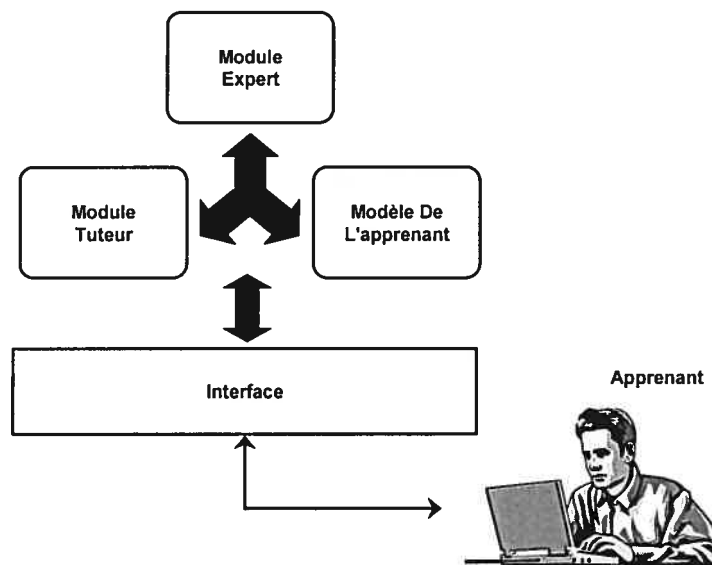


Figure 2-1 – Architecture classique d'un STI

2.2.1 Le module expert

Le module expert dans un STI, est le module chargé de fournir la connaissance à enseigner dans un domaine donné, indépendamment des aspects pédagogiques sur comment enseigner cette connaissance et indépendamment de la connaissance sur l'apprenant qui va suivre cette formation [Burns et Capps, 1988]. Il est constitué de l'expertise du domaine sous la forme de faits, de règles ou sous toute autre représentation qui peut faire l'objet d'un apprentissage ou qui peut servir dans le cadre de résolution de problèmes. Cette expertise du domaine est obtenue par consultation des experts qui ont à leur actif plusieurs années d'expérience dans le domaine en question [Anderson, 1988]. Le travail de recueillir et de codifier cette connaissance était et reste encore le travail le plus exigeant dans la réalisation d'un STI [Murray, 1999], et peut être accompli selon trois modèles : le modèle à boîte noire, le modèle à boîte transparente et le modèle cognitif :

- **Le modèle à boîte noire** : la modélisation du module expert selon ce principe essaye de trouver des manières de raisonner dans le domaine, qui n'exigent pas vraiment la codification de la connaissance qui est à la base de l'intelligence et du comportement

humain dans des contextes liés à ce domaine. Par exemple, un système peut utiliser une équation mathématique, qui produit par des traitements numériques ce que l'humain calcule par des traitements symboliques (comme dans le calcul d'une intégrale par exemple). Le système, vis-à-vis de certains paramètres en entrée, génère le bon comportement ou la bonne sortie, et de ce fait peut être utilisé pour juger l'exactitude du travail de l'apprenant. Ce principe a été utilisé dans des systèmes tels que SOPHIE pour l'entraînement aux diagnostics de pannes dans les équipements électroniques [Brown, Burton et Bell, 1975]. Le système même sans avoir une connaissance du courant électrique et de l'électronique similaire à celle de l'humain, pouvait utiliser ses modèles mathématiques et simuler le même résultat que l'humain. Seulement les réactions du système se limitaient à un jugement du genre "Ceci est juste" ou "Ceci est faux, recommencer". Pour articuler un peu plus les interactions de ce genre de systèmes, une technique, dite "*issue-based tutoring*", a été introduite dans le système WEST [Burton et Brown, 1976] et utilisée dans d'autres systèmes par la suite, comme dans Geometry Tutor [Anderson, Boyle, et Yost, 1985]. L'idée du "*issue-based tutoring*" est d'associer des patrons aux comportements de l'expert et aux comportements de l'apprenant et d'associer des explications à ces patrons. En plus de fournir un jugement sur l'exactitude du travail de l'apprenant, le système cherche à trouver dans le comportement de l'apprenant qui est jugé erroné ou non optimal, un *matching* avec l'un des patrons de comportement mémorisés. Quand la recherche de ce *matching* est fructueuse, le système présente l'explication associée au patron trouvé pour expliquer à l'apprenant les erreurs dans sa démarche.

- **Le modèle à boîte transparente :** la seconde façon de faire la codification du module expert est selon le principe de la *boîte transparente*, par l'utilisation de l'approche *système expert* comme dans le domaine de l'intelligence artificielle. Ceci implique de faire appel à un ingénieur de la connaissance et un expert du domaine qui vont définir les limites du problème, énumérer et formaliser les concepts clés du domaine, formuler un système pour mettre en œuvre cette connaissance, puis, et d'une manière itérative, tester et raffiner le système développé. L'idée d'utiliser les systèmes experts dans la création du module expert était très attrayante. Les systèmes experts sont caractérisés par la grande quantité d'informations sur le domaine qu'ils doivent avoir et par la nature même de cette information qui est articulée d'une manière assez similaire à la manière dont

l'humain articule la connaissance. De plus, les systèmes experts ont été utilisés avec succès dans plusieurs domaines comme la prédiction, le diagnostique, la planification ou le contrôle.

Cependant, les premières applications de cette approche dans la réalisation du module expert basé sur un système expert n'ont pas été concluantes. L'exemple classique est celui du module expert GUIDON [Clancey, 1982] qui était basé sur le système expert MYCIN [Shortliffe, 1976]. Son domaine d'application était le diagnostic médical des infections bactériologiques et il consistait en un ensemble de 450 règles du type *si-alors* comme celle indiquée dans la Figure 2-2. Bien que ces règles permettent au système de faire des diagnostics corrects, elles étaient très peu utilisables sur le plan pédagogique. Elles étaient trop condensées pour servir dans des explications à l'apprenant et le mécanisme de recherche par chaînage arrière ne correspondait pas aux mécanismes de raisonnement accomplis par l'humain.

IF

- The infection which requires therapy is meningitis
- Organisms were not seen in the stain of culture
- The type of infection is bacterial
- The patient does not have a head injury defect
- The age of the patient is between 15 and 55 years

Then

- The organisms the might have been causing the infection are
- Diplococcus pneumoniae(.75) and
 - Neisseria-meningitidis(.74)

Figure 2-2 – Exemple de règle dans MYCIN

- **Le modèle cognitif** : l'expérience de l'utilisation des systèmes experts dans GUIDON a démontré que pour créer des STIs vraiment efficaces, il fallait en plus de la codification de la connaissance de l'expert, penser à comment cette connaissance allait être déployée, car l'humain et le système expert n'utilisent pas cette connaissance de la même façon. Ceci a conduit au principe du *modèle cognitif* dans la réalisation des modules experts. Le but du modèle cognitif dans la réalisation des modules experts est la création d'une simulation de l'approche humaine dans la résolution de problèmes. [Anderson,

1988] a distingué trois types de connaissances qui peuvent constituer un modèle cognitif : les connaissances procédurales, les connaissances déclaratives et les connaissances causales.

La connaissance procédurale. Dans l'enseignement de domaines comme le calcul ou la résolution de problèmes, la connaissance principale à communiquer est procédurale. Elle montre comment les tâches et les actions sont effectuées, sous quelles conditions et dans quel ordre elles sont effectuées. Comme l'activité humaine dans ce contexte prend généralement la forme de cycles "*Reconnaître-Agir*", il était naturel de représenter l'aspect procédural par des systèmes de production avec des règles "*Si-Alors*". C'était le cas dans le système LISP Tutor [Reiser, Anderson et Farrell, 1985] et Geometry Tutor [Anderson, Boyle, et Yost, 1985].

La connaissance déclarative. À l'opposé de la connaissance procédurale, la connaissance déclarative sert à communiquer des ensembles de faits, *convenablement organisés* sur lesquels l'apprenant peut raisonner. Les domaines où elle s'applique sont par exemple la géographie et l'histoire. L'exemple classique de ce genre de systèmes est le système SCHOLAR de Carbonell [Carbonell, 1970]. Il était conçu pour l'enseignement de la géographie de l'Amérique du sud. Les différents faits et données liés à ce sujet ont été organisés dans un réseau sémantique qui semblait être très proche de la façon dont l'humain organise la connaissance. Le système utilisait ce réseau sémantique pour faire des inférences sur la matière à enseigner et sur le raisonnement de l'apprenant et permettre d'engager ce dernier dans des dialogues de type socratique.

La connaissance causale. Ce type de connaissance est relié à notre capacité à simuler mentalement et à raisonner sur des processus et phénomènes de nature dynamique. Elle prend généralement la forme de modèles qualitatifs et peut parfois être jumelée à des simulations qualitatives comme dans le cas du système SOPHIE [Brown, Burton et Bell, 1975].

Nous pensons que sur le plan pratique, réaliser un module expert, ne revient pas à choisir un modèle de représentation et utiliser le même modèle pour tout un cours. Un cours fait généralement intervenir plusieurs contextes d'apprentissage différents pouvant nécessiter des modèles de représentation différents. Ce qui fait qu'un module expert est un mélange de modes de représentation appropriés à chacun des contextes d'apprentissage. De plus, un module expert n'est pas formé uniquement de la représentation de la connaissance modélisée

sous une forme donnée, mais il est aussi un *système d'interprétation* de cette représentation [Polson, 1993].

2.2.2 Module tuteur

Le modèle tuteur est la composante d'un STI qui initie et gère les interactions pédagogiques avec l'apprenant [Half, 1988]. Dans certaines architectures il reçoit parfois le nom du module pédagogique ou du module d'instruction [Vasandani et Govindaraj, 1995]. Mais quel que soient les noms donnés, les responsabilités attribuées à cette composante restent les mêmes :

- = La première responsabilité du module tuteur est de faire *la sélection et le séquençement* du matériel pédagogique à présenter à l'apprenant. C'est la gestion du curriculum. Ce matériel pédagogique peut contenir des éléments comme des pages HTML présentant le contenu, des exercices de pratiques et des exercices d'évaluation. Ces éléments peuvent être classifiés par leurs natures : *déclarative* pour les éléments du contenu comme les pages HTML et *procédurale* pour les exercices pratiques et d'évaluation. Le séquençement de ces éléments doit respecter un certain nombre de principes en fonction de la nature de l'élément donné.

Le principe de proximité: donner la priorité aux concepts qui sont étroitement liés au contexte en cours et ainsi les présenter en premier.

Le principe de généralité: présenter les généralités d'un sujet avant de présenter les spécificités.

Le principe d'intégrité d'ordonnement: présenter les pré-requis d'un élément avant la présentation de celui-ci.

Des procédures de génération de curriculums très sophistiquées ont été introduites dans les STIs comme le système CREAM [Nkambou, 1996] et dans CREAM nouvelle version [Rouane, 1999]. Elles prenaient en compte les objectifs de formation à atteindre et le niveau initial du public ciblé pour cette formation. La génération de curriculum se faisant en deux étapes. Dans un premier temps un graphe de cours est généré (Figure 2-3) représentant les activités pédagogiques et les relations qui les organisent (principalement les relations de pré-requis). Puis une séquence de ces activités est créée à partir de ce graphe de cours en respectant les principes cités ci-dessus. Dans le système CREAM nouvelle version [Rouane, 1999], un mécanisme à deux niveaux de génération

de curriculum a été introduit. Le premier niveau faisait une génération initiale qui donnait une séquence destinée à un public cible relativement large (comme un groupe d'étudiants ou une classe), sans être destinée à un apprenant particulier. Le deuxième niveau faisait une génération adaptative ou ad hoc qui adaptait continuellement la séquence initiale au profil d'un apprenant particulier.

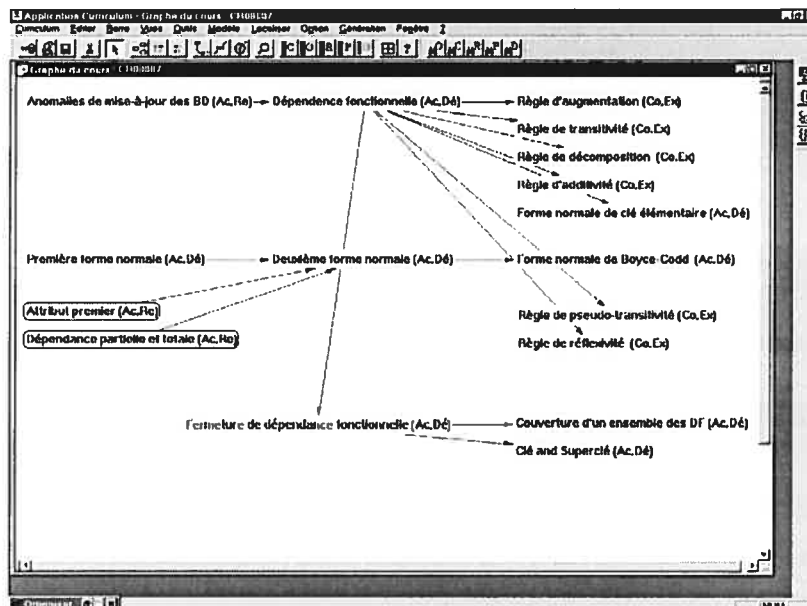


Figure 2-3 – Exemple de graphe de cours [Rouane, 1999]

- La deuxième responsabilité dont est chargé le module tuteur est d'apporter de l'aide à l'apprenant à la demande de ce dernier. L'aide demandée peut se rapporter à la matière enseignée (par exemple demander plus d'explications sur un concept particulier) ou à l'instruction (par exemple demander à voir sa progression ou la planification des activités à venir). En s'inspirant de l'activité du tuteur humain, [Half, 1988] suggérait que le module tuteur devait être capable de répondre aux questions que peut se poser un apprenant par rapport au sujet enseigné. Mais pragmatiquement, il faut reconnaître la difficulté d'une telle entreprise, surtout avec les difficultés que pose le traitement du langage naturel et la complexité de la modélisation nécessaire.
- La troisième responsabilité d'un module tuteur est d'apporter de l'aide à l'apprenant quand le système le juge nécessaire. Ce type d'aide entre généralement dans le cadre de résolution de problèmes ou dans l'exécution d'exercices. Ces interventions initiées par le systèmes peuvent être faites selon deux stratégies : le "Model Tracing" et le "Issue-based".

- **Stratégie avec “*Model Tracing*”** : c'est une stratégie très simple à mettre en œuvre. L'observation des actions de l'apprenant par le système se fait à la trace et le système intervient dès qu'une erreur est survenue [VanLehn, 88]. L'inconvénient de cette stratégie est que l'étudiant ne peut pas explorer des chemins erronés dans l'espace du problème et obtenir l'explication de ses erreurs en face de contradictions ou d'impasses. Cette stratégie a été mise en oeuvre dans des systèmes comme Steamer [Hollan et al., 1984] sur le fonctionnement d'une turbine à vapeur sur un navire, et le système Quest [White et Frederiksen, 1985] sur la détection de pannes électriques.
- **Stratégie avec “*Issue-based*”** : c'est une stratégie plus élaborée que la précédente. Connue aussi sous le nom de *possibilités et exemples*, elle est basée sur le principe d'exploitation d'opportunités. La démarche consiste à préparer d'avance pour le système, des opportunités possibles pour l'apprenant dans l'étape d'exécution, et qui sont considérées comme étant les opportunités représentant le mieux la démarche experte. Ces opportunités sont constituées généralement de suites d'actions sous la forme de patrons (patterns) de comportements. Le système intervient alors dans les cas où l'apprenant fait une suite d'actions qui s'apparentent à l'une des opportunités mémorisées, même si elle est jugée correcte, mais qui n'appartient pas à ensemble des démarches expertes. Le système utilise les explications et les exemples associés à chaque opportunité pour expliquer à l'étudiant les faiblesses de sa démarche pour pouvoir l'améliorer dans le futur. Cette stratégie a été exploitée dans WEST [Burton, 1982].

La Figure 2-4 illustre la différence entre les réactions d'un STI en présence d'une stratégie “*Model Tracing*” et d'une stratégie “*Issue based*”.

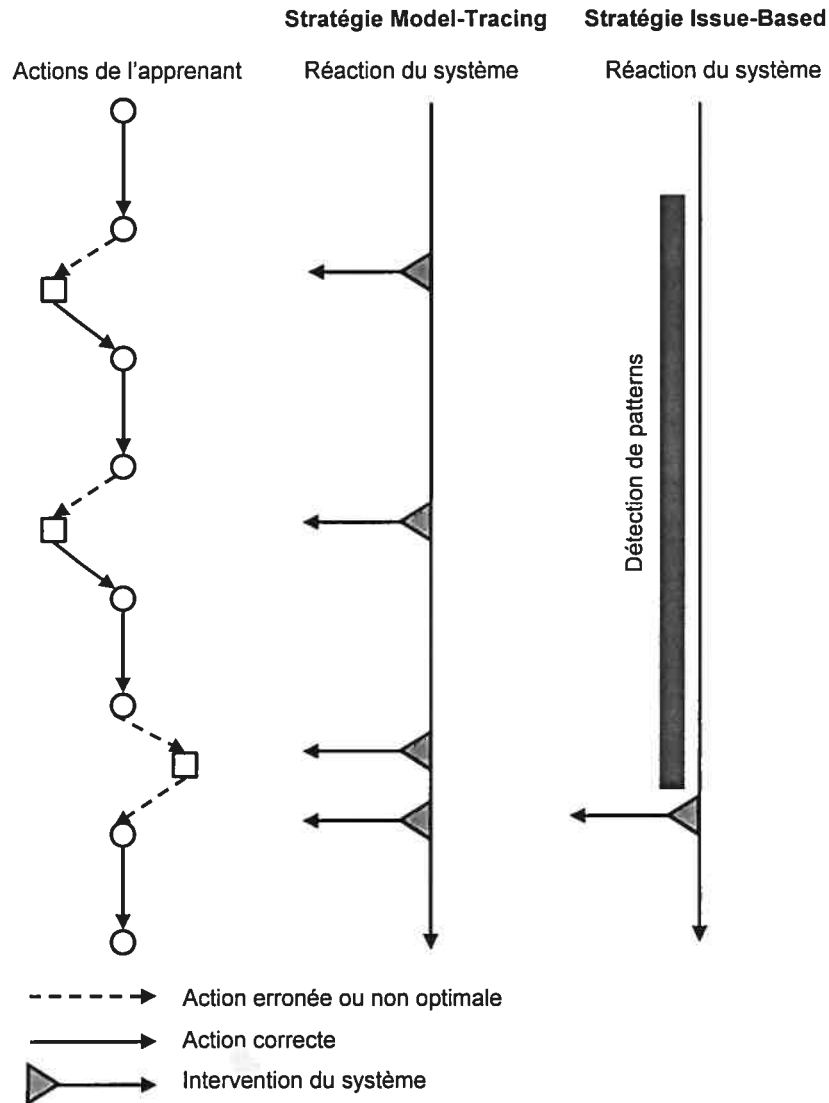


Figure 2-4 – Réactions du système dans les stratégies Model Tracing et Issue-based

2.2.3 Modèle de l'apprenant

La modélisation de l'utilisateur (ou le “*User modeling*”) est un élément nécessaire pour la réalisation d'applications adaptatives dans un grand nombre de domaines informatiques comme la présentation de l'information personnalisée, l'adaptation des interfaces usagers, la génération d'aide personnalisée [Jameson, 1999]. Les chercheurs dans les STIs étaient parmi les premiers à reconnaître le besoin de modéliser l'utilisateur, dans ce cas l'apprenant, dans les systèmes didactiques et fournir ainsi le moyen de permettre un comportement *adaptatif* de

ces systèmes [Self, 1974]. La modélisation individuelle de l'apprenant, à laquelle on réfère généralement par *le modèle de l'apprenant* (ou "*Student model*"), est une caractéristique distinctive des systèmes tutoriels intelligents par rapport à l'enseignement assisté par ordinateur conventionnel.

Dans un sens strict, le modèle de l'apprenant est la représentation de la connaissance de l'apprenant dans un STI [VanLehn, 1988]. Mais dans un sens plus large, il est défini comme étant toute composante d'un STI ou représentation de connaissance qui permet à un STI d'assister un apprenant d'une façon individualisée et personnalisée [Self, 1999]. Grâce à ce modèle, les systèmes STIs sont capables de suivre et d'analyser les besoins individuels de l'apprenant et fournir le feedback approprié et les ajustements du matériel éducatif pour faire face et traiter les différentes difficultés pédagogiques susceptibles de survenir dans une situation d'apprentissage. [Self, 1988] a identifié vingt différentes utilisations du modèle de l'apprenant dans une situation d'apprentissage. Après analyse, il a noté que toutes ces fonctions peuvent être classées en six catégories de fonctions :

- *Les fonctions correctives* : pour aider à éliminer les erreurs dans la connaissance de l'apprenant.
- *Les fonctions élaboratives* : pour aider à corriger une connaissance incomplète de l'apprenant.
- *Les fonctions stratégiques* : pour aider à initier un changement important dans la stratégie pédagogique suivie, quand les deux premiers types de fonction ne suffisent plus.
- *Les fonctions de diagnostic* : pour aider à diagnostiquer les erreurs dans la connaissance de l'apprenant.
- *Les fonctions prédictives* : pour aider à prédire les réactions et réponses probables de l'apprenant (fonction utile pour la planification).
- *Les fonctions évaluatives* : pour aider à évaluer l'apprenant.

La modélisation du modèle de l'apprenant est caractérisée par sa connexion avec plusieurs disciplines scientifiques telles les sciences de l'éducation, la psychologie, les sciences sociales et les sciences cognitives. Idéalement, le modèle de l'apprenant doit inclure tous les aspects du comportement et de la connaissance de l'apprenant pouvant avoir des répercussions sur

sa performance et son apprentissage. Plusieurs types de modèle d'apprenant ont été proposés [Zhou et Evens, 1999]. Certains modèles, dits modèles de recouvrement, tentent de représenter la connaissance de l'apprenant par projection sur la connaissance du domaine [Ragnemalm, 1995], d'autres s'intéressent aux violations de contraintes commises par l'apprenant, ce sont les modèles basés sur les contraintes [Ohlsson, 1992], alors que d'autres analysent les aspects émotionnels et affectifs de l'apprenant [Frasson et Abou-Jaoude, 1998].

Le modèle de l'apprenant basé sur le recouvrement

Le modèle de recouvrement ou "*the overlay model*" [Ragnemalm, 1995] [Nkambou, 1996] [Rouane, 1999] est le type de modèle de l'apprenant le plus utilisé dans les STIs. Il tente de représenter la connaissance de l'apprenant par superposition (projection) à celle de l'expert du domaine. Il tente de déterminer ce que connaît l'apprenant et à quel niveau, et ensuite de déduire ce qui lui reste à connaître et ce qu'il doit mieux connaître. Ceci permet au système de planifier les activités pédagogiques pour combler les différences observées.

Le modèle de recouvrement a été dérivé en plusieurs variantes : le modèle de recouvrement simple, le modèle de recouvrement avec erreurs, le modèle de perturbation et le modèle de perturbation avancé (Figure 2-5). Ces dérivations ont été effectuées en fonction de trois types de connaissance : la connaissance du domaine (ou l'expertise du domaine), la connaissance erronée (les erreurs fréquentes) et la connaissance propre de l'apprenant (qui est la connaissance introduite par l'apprenant et dont le système ne peut apporter aucun jugement, ni positif, ni négatif).

Dans le modèle de recouvrement simple, un seul type de connaissance est représenté, la connaissance du domaine. La connaissance de l'apprenant est modélisée comme un sous-ensemble de la connaissance du domaine. La connaissance erronée et la connaissance propre de l'apprenant ne sont pas prises en compte (A, Figure 2-5).

Dans le modèle de recouvrement avec erreurs, la connaissance du domaine et la connaissance erronée sont toutes deux prises en compte. La connaissance de l'apprenant est modélisée comme un sous-ensemble de la connaissance du domaine et de la connaissance erronée. La connaissance propre de l'apprenant n'est pas considérée (B, Figure 2-5).

Le modèle de perturbation et le modèle de perturbation avancé sont très similaires, respectivement, au modèle de recouvrement simple et au modèle de recouvrement avec erreurs, mais avec prise en compte de la connaissance propre de l'apprenant (C et D, Figure 2-5).

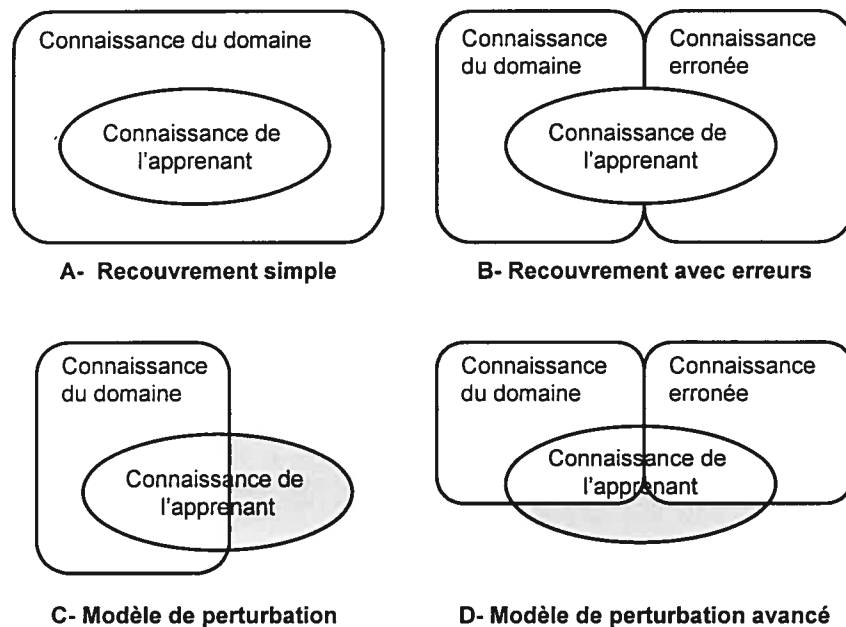


Figure 2-5 – Quatre vues de relations entre les connaissances de l'apprenant et celles du domaine

Le modèle de l'apprenant basé sur les contraintes

Le modèle de l'apprenant basé sur les contraintes (*The constraint-based model ou CBM*) proposé par Ohlsson [Ohlsson, 1994] représente les deux connaissances, celle de l'apprenant et celle du domaine sous la forme de contraintes, où les contraintes représentent les principes de base liés au domaine enseigné. Une contrainte est caractérisée par une clause de pertinence et une clause de satisfaction. La clause de pertinence est une condition qui doit être satisfaite avant que la contrainte ne soit considérée dans un contexte d'une solution donnée. Une fois la clause de pertinence satisfaite, la clause de satisfaction doit aussi être satisfaite pour que la solution soit considérée comme correcte.

Les contraintes qui forment le domaine de connaissance peuvent être considérées comme un ensemble de règles, où chaque contrainte est une règle sous la forme “Si (clause de pertinence) est vraie, Alors (clause de satisfaction) doit aussi être vraie.” Alternativement, cet ensemble peut être considéré comme une librairie d’erreurs (*Bug library*). Dans cette vision, une contrainte peut être représentée comme suit : “Quand une (clause de satisfaction) n’est pas vérifié pour une (clause de pertinence), on considère que l’apprenant a commis une erreur.”

Le modèle de l’apprenant basé sur les contraintes est un ensemble de contraintes qui n’ont pas été respectées par l’apprenant et qui représente la connaissance du domaine où l’apprenant éprouve des difficultés. Ohlsson indique qu’il est dangereux de considérer que l’ensemble des contraintes que l’apprenant a respectées, représente ce que connaît l’apprenant car il pourrait avoir satisfait ces contraintes par inadvertance et sans compréhension des concepts impliqués. Cependant, la violation d’une contrainte indiquerait toujours une inconsistance dans la compréhension de l’apprenant.

La complexité de la mise en œuvre du modèle de l’apprenant

La modélisation de l’apprenant est une tâche très difficile qui exige des algorithmes de diagnostic très complexes. [Self, 1990] a indiqué que le diagnostic du profil de l’apprenant doit être fait avec l’objectif constant de l’utilité (computationnelle) plutôt que de chercher la fidélité cognitive :

“Student modeling is not about building exact cognitive models. If it were, we would have to solve all the problems of cognitive science, and teach a machine to be a cognitive scientist, before we could build a student model. We only need to model the student to the level of detail necessary for the teaching decisions we are able to take.” [Self, 1990]

La prise en considération de ces aspects pratiques et pragmatiques a conduit les chercheurs dans ce domaine à explorer de nouvelles voies pour la modélisation de l’apprenant. Certaines voies font intervenir le tuteur humain pour contourner ce problème de complexité alors que d’autres font intervenir l’apprenant directement comme dans le cas du *modèle ouvert de l’apprenant* [Dimitrova, Self et Brna, 1999].

Une autre source de difficulté et qui reste toujours présente, est l’initialisation du modèle de l’apprenant. Quel que soit le modèle de représentation choisi, le modèle de l’apprenant doit

être initialisé au premier contact de l'apprenant avec le système. [Barr et Feigenbaum, 1982] ont identifié quatre façons de le faire qui sont toujours valables :

- *Implicitement*, par supervision du comportement de l'apprenant dans sa résolution de problème.
- *Explicitement*, par utilisation de dialogue direct avec l'apprenant.
- *Structurellement*, en reflétant les difficultés inhérentes au domaine.
- *Historiquement*, en faisant des hypothèses en fonction des expériences précédentes de l'apprenant.

2.2.4 Interface de communication avec l'apprenant

L'interface de communication avec l'apprenant est la composante du STI qui contrôle la communication et les interactions entre l'apprenant et le système. Elle est chargée de faire la traduction dans les deux sens, entre une représentation interne du système et une présentation externe de l'interface qui soit ergonomique et compréhensible par l'apprenant. Cette composante est extrêmement importante, car de sa qualité peut dépendre le succès ou l'échec d'un STI, et ceci indépendamment du degré "d'intelligence" et de sophistication incorporés dans les composantes internes non visibles du système. En fait, ce serait une erreur de la considérer comme une composante secondaire du STI pour deux raisons.

Premièrement, quand un STI tente de communiquer avec l'apprenant pour, par exemple, lui présenter du contenu ou une explication, l'interface du système peut mettre en valeur ou affecter négativement la présentation. Les qualités d'une interface facile d'utilisation et attrayante pour l'apprenant sur le plan visuel, sont des éléments cruciaux qui peuvent influencer l'acceptation de tout le système par l'apprenant.

Deuxièmement, les progrès dans les technologies du multimédia et de la réalité virtuelle offrent des outils et des méthodes de communication de plus en plus sophistiqués qui font reculer les limites de l'utilisation de l'ordinateur dans l'enseignement.

Prenons le cas des habiletés motrices par exemple. Ce sont des habiletés très différentes des habiletés intellectuelles et qui visent l'apprentissage de mouvements coordonnés. Ces habiletés étaient délibérément omises dans les STIs conventionnels car leurs interfaces n'étaient pas en mesure de les supporter. Mais les avancées dans le domaine de la réalité

virtuelle et celui des instruments de simulation mécaniques (e.g. les souris 3D à retour de force) permettent aujourd'hui d'étendre les possibilités des interfaces de communication dans les STIs au-delà du cadre conventionnel "clavier-souris-écran".

2.3 Difficultés et critiques des approches STI actuelles

La recherche dans le domaine des STIs s'est traditionnellement focalisée sur les domaines de connaissances simples reliées à l'apprentissage d'habiletés élémentaires dans des disciplines comme les mathématiques, l'électricité ou l'électronique, et les langages de programmation. Ces disciplines scientifiques de nature formelle et fermée se prêtaient plus facilement à l'automatisation que d'autres disciplines scientifiques moins formelles et plus ouvertes, qui parfois ont reçu le nom de "sciences molles" ou *soft-science*, telles les sciences humaines et sciences sociales. L'expérience de recherche acquise dans les ITSs sur les domaines fermés, était très difficilement généralisable à ce genre de disciplines. Ceci est dû principalement à des difficultés sur le plan de la théorie, sur le plan de la modélisation et sur le plan de la pédagogie.

2.3.1 Difficultés sur le plan de la théorie

La première source de difficulté dans les STIs vient des bases théoriques utilisées dans leurs développements. Les concepteurs et les chercheurs dans ce domaine ont tenté depuis le début d'adopter et de créer des théories pour guider la conception de ces systèmes [Wenger, 1987] [Polson et Richardson, 1988]. Seulement ces théories, entièrement *discursives* [Self, 1990], n'ont fourni aucune sorte de formalisme à leurs contenus et sont restées sur le plan des principes. Or la construction des STIs est aussi un travail d'ingénierie, puisque le but est de construire un système informatique qui fonctionne effectivement, et exige donc plus de précision et de formalisme dans les principes et théories utilisées. De plus les situations éducatives sont très complexes et on ne peut connaître d'une façon parfaite toutes leurs composantes [Bru, 2001].

Si la conception en général est considérée comme un mélange d'art et de science, le manque de formalisme dans les théories de base des STIs a rendu leur conception presque exclusivement un art [Self, 1999]. La conception d'un STIs n'est pas plus complexe que celle d'un avion par exemple. Seulement la conception d'un avion se base elle sur des théories

ayant des bases formelles et précises, comme l'aérodynamique et les sciences des matériaux, alors que la conception des STIs ne repose pas sur des théories aussi formelles.

Devant ce constat, les concepteurs de STIs ont essayé de trouver des alternatives à l'absence d'une théorie pure des STIs, en essayant d'adopter des théories annexes comme les théories d'apprentissage ou les théories d'intelligence artificielle. Mais leurs principes ne sont pas directement applicables dans un contexte STI. Clancey [Clancey, 1986] soutient que la conception des STIs n'est pas une affaire de simple mise en pratique de principes bien connus en Intelligence Artificielle (ou en théorie d'apprentissage).

Murray [Murray, 1996-a et b], ajoutait que les techniques avancées en IA comme la planification, le raisonnement à base de cas, la compréhension du langage naturel, les réseaux de neurones, la logique floue, ou la résolution de problème à base de règle doivent être utilisées avec une grande prudence et modération si nous voulons produire effectivement des STIs utilisables, car l'introduction de telles techniques augmente considérablement la complexité des outils de conception et la tâche des concepteurs alors que les bénéfices que l'on souhaite obtenir par l'utilisation de telles techniques ne sont pas certains [Du Boulay, 2000]. [Cerri, 1999] va dans le même sens et prône la simplicité comme élément essentiel dans la conception de ces systèmes :

“Within the Artificial Intelligence and Education community, we notice a major concern that reaches the same conclusion, i.e. complexity in AI & ED research makes it almost unviable for AI & ED real applications.... Cognitive simplicity in the conception and design of new applications becomes a must for any concrete dissemination of research results as well as for most applicative efforts.” [Cerri, 1999].

D'autre part, les disciplines les plus influentes sur l'efficacité d'utilisation des STIs, telles les sciences cognitives et les sciences d'apprentissage, se prêtent très difficilement à la modélisation et l'automatisation. De ce fait, les STIs qui essaient de se baser principalement sur ces théories se trouvent alors extrêmement limités. Considérons par exemple, la définition de Bloom pour les objectifs pédagogiques [Bloom, 1969] [Bloom, 1978] qui est très utilisée dans l'enseignement conventionnel et adoptée par certains STIs [Nkambou, 1996] [Rouane, 1998]. Selon cette définition, un objectif pédagogique est relatif à une habileté à acquérir par l'apprenant, comme l'habileté de *reconnaître* ou l'habileté d'*énumérer*. Il est clair que pour ce genre d'habileté, les concepteurs des STIs peuvent trouver une façon

d'informatiser l'atteinte et le contrôle de ces habiletés. Mais ce n'est pas le cas pour toutes les autres habiletés définies dans le même modèle Bloom, comme les habiletés de *discuter* ou d'*expliquer*.

2.3.2 Difficultés dans la modélisation du domaine

La modélisation du contenu des STIs a été très marquée par *l'hypothèse de la représentation symbolique*, suivant les travaux de Newell, A. et Simon, H. en Intelligence Artificielle [Luger, 1998]. Anderson [Anderson, 1988] considérait que par définition, un STI ne peut être construit que pour des domaines pour lesquels il existe un système expert.

Les concepteurs assumaient alors que la connaissance pouvait exister d'une façon *objective* et indépendante dans le monde externe [Self, 1990] et qu'on pouvait donc la décrire en terme de faits, de règles et de principes, représentés symboliquement et hiérarchiquement, et qu'on pouvait la faire acquérir par un apprenant d'une façon incrémentale [Self, 1999]. Par conséquent, ils ont investi leurs efforts dans le développement de représentations généralement complexes de cette connaissance. Cette façon de faire dans la représentation de la connaissance aurait probablement fonctionné si la connaissance existait d'une manière claire, objective et bien délimitée. Mais la réalité est tout autre comme l'expliquait [Murray, 1996-b] :

“Human knowledge does not exist in neatly defined, clearly named packages---it is inherently complex, densely connected, fuzzy, and ambiguous. Yet to use knowledge in AI systems we try to represent it in individual units with clear structure. The tension between the organic nature of knowledge and our need to modularize it leads to a number of issues for ITS knowledge representation.”

La représentation de la connaissance à la manière de l'IA dans les STIs, c'est à dire en décortiquant la connaissance en ses constituants élémentaires, avait plusieurs inconvénients. Le plus important de ces inconvénients, en dehors de la complexité évidente dans la conception, est l'atomisation de la connaissance et la perte du discours narratif dans la matière à enseigner.

La complexité dans la conception de ces systèmes est due essentiellement à la complexité dans l'acquisition de connaissances, comme c'est le cas dans tout système basé sur le principe du système expert. Luger [Luger, 1998] soutient que cette connaissance est souvent

inaccessible d'une façon consciente à l'expert du domaine, et encore moins à l'ingénieur de connaissance. Après des années de pratique, l'expertise devient un tout intégré opérant largement en dehors de la conscience, et par conséquent très difficile à décrire par l'expert, qui est capable de démontrer des compétences et des aptitudes de haut niveau, mais sans être capable d'expliquer en détail la mécanique qui les régit. Cette limitation est intrinsèque au paradigme du Système Expert et devient donc une limitation intrinsèque aux STIs qui utilisent ce paradigme [Patel et Kinshuk, 1997].

Le processus de modélisation selon cette approche, lorsqu'il est accompli, aboutit généralement à une représentation hiérarchique du contenu, qui subdivise la connaissance en ce qui peut être considéré comme étant les constituants *atomiques*¹ de cette connaissance [Rodenburg, 2001]. Dans les STIs basés sur l'approche *CREAM* par exemple [Rouane, 1998] [Nkambou, 1996], qui adoptent cette décomposition, la connaissance du domaine est subdivisée en une hiérarchie de concepts élémentaires ayant chacun une définition propre, des exemples, des contres-exemples et étant relié aux autres concepts (dans une structure en réseau dite: *le modèle des concepts*) par différents liens de généralisation/spécialisation, d'analogie, de déviation et autres.

Mais arrivé à l'étape de l'enseignement effectif, le STI qui s'appuie sur une telle représentation de connaissance n'est pas capable de donner un vrai enseignement. La myriade de petits concepts à enseigner à l'apprenant fait disparaître le sens global des idées. Rodenburg [Rodenburg, 2001], exprime très clairement cette problématique:

“...If we're talking about instruction -- by which I mean the integration of information into a cohesive and meaningful whole -- then I think that instructional design demands a more sophisticated approach. Meaningful learning, even in a training context, is not about absorbing small "info-bytes", but about struggling to understand, at a conceptual or holistic level, the underlying context and principles that relate that information to a larger picture.”

Un autre inconvénient qui est lié à la décomposition de la connaissance, est l'absence du discours narratif dans la phase de présentation. Présenter la connaissance à l'apprenant, sous format d'une suite de paquets indépendants n'est pas naturel et va à l'encontre de notre

¹ Atomic "bits"

habitude à apprendre dans un discours [Weller, 2000]. [Murray, 1996-a] pense que la connaissance qui relie un groupe de paquets doit être explicitement représentée car elle met en évidence la connaissance émergente de la compréhension de ce groupe de paquets. Lesgold [Lesgold, 1988] qualifiait cette connaissance de “colle” de curriculum. Notons que les discours narratifs contenus dans les livres éducatifs¹ contiennent déjà cette connaissance et d’une façon implicite.

Fisher [Fisher, 1987] suggère que le discours est important à la compréhension, car l’encadrement des événements dans une histoire² permet aux individus d’interpréter leurs environnements et ainsi mieux raisonner sur le monde. Ces environnements qui constituent des contextes³ pour l’apprentissage sont très importants [Gagné, 1985], car le rappel des éléments mémorisés exige souvent le rappel du contexte dans lequel ces éléments ont été appris [Luger, 1994].

2.3.3 Difficultés dans la modélisation de l’apprenant

Le modèle de l’apprenant, qui est la représentation de la connaissance de l’apprenant dans un STI [VanLehn, 1988], est très important pour l’adaptation et l’individualisation de la formation. Mais si les principes derrière ce modèle sont faciles à comprendre, ils sont par contre extrêmement difficiles à mettre en pratique et ce pour plusieurs raisons.

- Premièrement, ce modèle est complètement dépendant du modèle de la connaissance du domaine, et par conséquent, il est fortement affecté par toutes les difficultés rencontrées dans la construction du modèle de domaine abordées dans le sous-chapitre précédent.
- Deuxièmement, ce modèle est très difficile à initialiser et à maintenir. On est limité parfois à l’avis subjectif de l’apprenant pour déterminer les niveaux de maîtrise de la connaissance du domaine.

¹ par exemple les livres de référence largement utilisés dans la formation supérieure

² Composante du discours

³ plus ils sont variés, meilleur est l’apprentissage

- Troisièmement, les niveaux de maîtrise sont très difficiles à inférer et on s'appuie généralement sur les performances obtenues dans les exercices pour déterminer ces niveaux de maîtrise. Or dans la pratique, il est très rare qu'un exercice ne concerne qu'une seule connaissance, et encore moins un niveau de maîtrise précis de cette connaissance. De plus l'équivalence entre performance et maîtrise de connaissance, n'est pas toujours évidente.

Le modèle de l'apprenant reste l'un des problèmes complexes dans le domaine des STIs [Du Boulay, 2000]. Il est nécessaire pour l'individualisation de la formation, mais les principes retenus pour sa construction le rendent impraticable:

“...It is debatable whether the cost of constructing very detailed, complex user models that are runnable and have to be maintained all the time is worthwhile in terms of the gain in teaching efficiency.” [Self, 1990-a]

2.3.4 Difficultés dans la pédagogie

Une étude rapide de l'activité d'enseignement qu'accomplit un STI révèle l'existence d'un cycle répétitif d'opérations (Figure 2-6) qui d'une façon générale, peut être ramené à trois étapes [Rouane, Frasson et Kaltenbach, 2002-a]:

- Présentation du contenu non-évaluable
- Présentation du contenu évaluable
- Prise de décisions pédagogiques

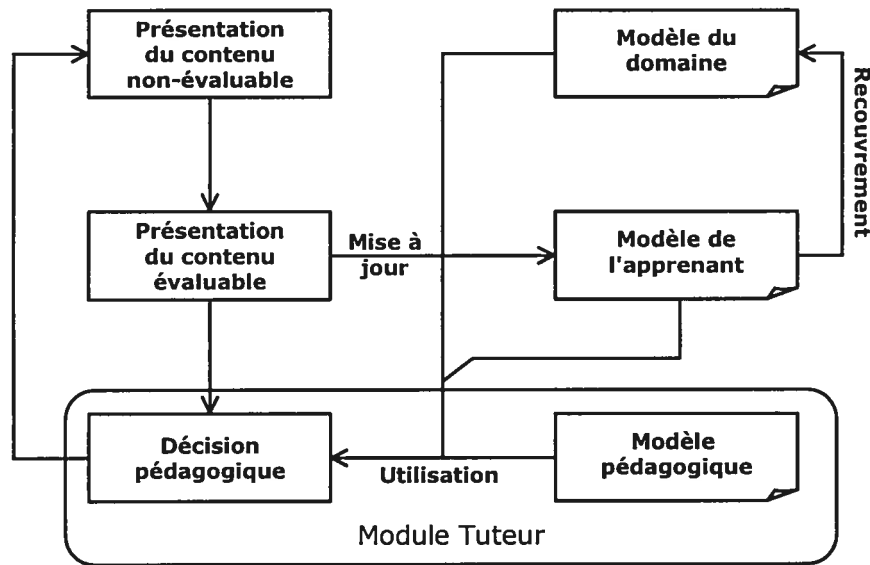


Figure 2-6 – Cycle des opérations dans l'activité d'enseignement

Le contenu non-évaluable est le contenu qui n'est pas sujet à une évaluation du degré et de la nature de l'acquisition de connaissances que peut accomplir un apprenant qui utilise ce contenu (par exemple un texte annexe à lire). De ce fait, le contenu non-évaluable ne modifie pas le modèle cognitif de l'apprenant et donc il n'intervient pas dans la prise des décisions pédagogiques faites par le modèle du tuteur.

Il faut faire la distinction entre un contenu non-évaluable et un contenu non-intelligent (ou dumb). Un contenu peut être intelligent et considéré comme non-évaluable s'il n'a pas la capacité de juger les performances de l'apprenant qui l'utilise ou s'il n'a pas la capacité de communiquer son jugement au module tuteur. Et par conséquent il n'influence pas les décisions pédagogiques de ce dernier.

Un fait très important à noter est que les contenus non-évaluables constituent la plus grande partie du matériel pédagogique mis à la disposition de l'apprenant et que la connaissance du domaine est dans la majorité des cas présentée à l'apprenant via des contenus non-évaluables, tels le texte, le schéma, l'image, la vidéo, etc.

Malgré cette importance, il reste que l'un des grands défauts des STIs actuels est le fait qu'ils présentent ce type de contenu à l'apprenant d'une façon très passive, alors que c'est à ce

niveau que l'acquisition des connaissances est faite par l'apprenant et c'est à ce niveau aussi que des erreurs de compréhension peuvent être commises par celui-ci.

Ces erreurs de compréhension ont des conséquences très graves lorsqu'elles sont commises et non corrigées à ce niveau :

1. Elles sont susceptibles de provoquer une chaîne d'autres erreurs de compréhension en cascade, qui ont pour conséquence d'augmenter considérablement la charge cognitive de l'apprenant qui se trouve alors de moins en moins motivé et peut être découragé à poursuivre l'exploration du reste du contenu. L'analyse et la correction de ces erreurs deviennent de plus en plus difficiles au fur et à mesure que cette chaîne d'erreurs devient de plus en plus longue.
2. Elles changent la nature de l'utilisation du contenu évaluable, comme les exercices, qui deviennent dans ce contexte des outils de détection d'erreurs de compréhension et non des outils d'évaluation. Or, ces ressources sont très précieuses si nous les comparons à du contenu non-évaluable, comme le texte par exemple: elles sont très difficiles à concevoir d'une part et d'autre part elles ne peuvent pas être présentées plus d'une fois.

2.4 Conclusion

Trouver des solutions pragmatiques et réalistes à ces difficultés n'est pas d'ordre technologique. Ce n'est pas la seule introduction de technologies comme le XML ou les systèmes multi-agents qui va nous permettre d'y arriver. Nous pensons que la solution commence par un examen en détail : 1) du processus de conception pédagogique qui mène à la création de cours, 2) des mécanismes d'apprentissage et des facteurs qui les influencent, et 3) de la nature de la connaissance relative au domaine enseigné. La discussion de ces points fait l'objet du prochain chapitre.

CHAPITRE 3

LA CONCEPTION PÉDAGOGIQUE ET LES MÉCANISMES D'APPRENTISSAGE

La conception pédagogique (*instructional design*) est une discipline scientifique concernée par la recherche sur les théories et les stratégies pédagogiques et par les processus de développement et d'implémentation de ces théories et de ces stratégies [ARL, 2004]. Dans une perspective pratique, la conception pédagogique est le processus de développement systématique de spécifications pédagogiques en tenant compte des théories et de mécanismes d'apprentissage. Ce processus englobe généralement les phases d'analyse des besoins de formation et des objectifs pédagogiques, de développement, d'implémentation, d'évaluation et de maintenance du matériel pédagogique.

3.1 La conception pédagogique : du behaviorisme au constructivisme

3.1.1 L'approche behavioriste

La conception pédagogique s'est développée dans les années 1960s, une époque où le paradigme behavioriste prédominait en psychologie générale, et en psychologie pédagogique (*The educational psychology*) en particulier.

Le paradigme behavioriste présume que les seules choses “vraies” en psychologie (ou qui méritent d’être étudiées) sont les choses qui peuvent être vues et observées. Les aspects mentaux d’une personne, conscients ou inconscients, ne peuvent être observés, mais nous pouvons observer comment cette personne agit, se conduit, se comporte. Ainsi, le behaviorisme insiste sur ce que *fait* une personne, et non ce que *pense* ou ce que *ressent* cette personne. Par conséquent, un apprentissage basé sur une théorie behavioriste focalise uniquement sur le comportement de l’apprenant et sur les facteurs et les conditions qui peuvent l’aider à acquérir un comportement jugé satisfaisant. Le behaviorisme considère que l’apprentissage n’est que l’acquisition de nouveaux comportements basés sur le conditionnement.

On retrouve les mêmes racines behavioristes tant dans la conception pédagogique que dans la psychologie pédagogique. Les chercheurs qui ont marqué cette discipline de la conception pédagogique, tel que Gagné, Merrill et Briggs, travaillaient tous dans la psychologie pédagogique. Mais avec la révolution cognitive des années 1970s, la psychologie pédagogique s’est différenciée de la conception pédagogique pour s’approcher davantage des sciences cognitives [Resnick, 1981] :

“An interesting thing has happened to instructional psychology. It has become part of the mainstream of research on human cognition, learning, and development. ... it is now difficult to draw a clear line between instructional psychology and the main body of basic research on complex cognitive processes. Instructional psychology is no longer basic psychology applied to education. It is fundamental research on the processes of instruction and learning.” [Resnick, 1981].

Cet état de fait a laissé la conception pédagogique relativement isolée durant les années 1980s. La nouvelle psychologie pédagogique considérait que la conception pédagogique et ses théories était trop *ad hoc*, s’intéressant beaucoup plus à la conception de stimulus qu’aux processus cognitifs eux-mêmes [Wilson et Cole, 1996]. De l’autre côté, la littérature en conception pédagogique a souvent ignoré le développement des théories cognitives, ce qui a résulté en des théories de l’éducation en dehors des principaux courants de la psychologie cognitive.

Ce n’est que vers la fin des années 1980s et le début des années 1990s, que la conception pédagogique et la psychologie pédagogique ont commencé à se concerter et ceci grâce à leur

intérêt commun pour les sciences cognitives et le cognitivisme. La Table 3-i résume l'historique de la relation entre la conception pédagogique et la psychologie pédagogique.

| Période | 1960-1975 | 1976-1988 | 1989 - présent |
|--|--|--|--|
| Paradigme dominant | Behaviorisme | Psychologie du traitement de l'information | Construction de la connaissance / Médiation sociale |
| Statut de la conception pédagogique | Conception pédagogique émergente | Conception pédagogique engagée dans la théorie / développement des modèles | Conception pédagogique engagée dans un processus de redéfinition |
| Statut de la psychologie éducative | Béhaviorisme | S'approche des nouvelles tendances dans les sciences cognitives | S'approche des nouvelles tendances constructiviste |
| Relation entre la conception pédagogique et la psychologie éducative | Les deux disciplines sont étroitement alignées | Les deux disciplines sont en divergence | Les deux disciplines sont en dialogue |

Table 3-i – Histoire de la relation entre la conception pédagogique et la psychologie éducative [Wilson et Cole, 1996]

3.1.2 L'approche cognitive

Le cognitivisme est une approche psychologique qui s'intéresse aux différentes activités mentales que peut effectuer un individu telles que : la réflexion, la mémorisation, le rappel, ou l'utilisation du langage. Bien que la psychologie cognitive ait émergé à la fin des années 1950s, ce n'est qu'à la fin des années 1970s qu'elle a commencé à réellement influencer le domaine de la conception pédagogique.

L'influence des sciences cognitives sur la conception pédagogique a déplacé le centre d'intérêt des pratiques behavioristes, qui insistaient principalement sur le comportement extérieur, aux pratiques cognitivistes qui insistent plutôt sur les processus mentaux internes et sur comment ils peuvent être utilisés pour promouvoir un apprentissage efficace. Les modèles de conceptions pédagogiques développés dans la tradition behavioriste, comme l'analyse des tâches par exemple, n'ont pas été complètement rejetés, mais plutôt remodelés

dans de nouveaux modèles qui prennent en compte des aspects cognitifs de l'apprentissage, comme la codification de la connaissance, la mémorisation de l'information, le rappel de cette information, et l'intégration de connaissances nouvelles avec la connaissance déjà acquise [Saettler, 1990]. Certains aspects behavioristes peuvent être acceptés dans un contexte cognitiviste, mais l'importance qu'on leur accorde et leur interprétation ne sont pas les mêmes :

“Cognitive theorists recognize that much learning involves associations established through contiguity and repetition. They also acknowledge the importance of reinforcement, although they stress its role in providing feedback about the correctness of responses over its role as a motivator. However, even while accepting such behavioristic concepts, cognitive theorists view learning as involving the acquisition or reorganization of the cognitive structures through which humans process and store information.” [Good et Brophy, 1990]

La conception pédagogique a repris beaucoup d'idées cognitivistes qui ont grandement influencé la façon dont le matériel pédagogique est conçu et présenté à l'apprenant. Parmi ces idées cognitivistes on retrouve les notions suivantes [Mergel, 1998] :

- **La notion du modèle psychologique du traitement de l'information à trois étapes** : quand une nouvelle information est présentée à un apprenant, cette information est traitée cognitivement en trois étapes : 1) elle passe en premier par le registre sensoriel, 2) puis elle est traitée dans la mémoire à court terme, 3) et enfin elle est transférée dans la mémoire à long terme pour son stockage permanent.
 - **Le registre sensoriel** : ce système reçoit en permanence des flux d'informations des différents sens. Ces flux peuvent durer de moins d'une seconde à quatre secondes et généralement disparaissent après par désagrégement ou par remplacement. Bien que toutes les informations véhiculées par ces flux soient suivies et contrôlées par le cerveau de l'individu, la majorité de ces informations n'arrive jamais à la mémoire à court terme et seules celles jugées importantes y parviennent.
 - **La mémoire à court terme** : les informations sensorielles en entrée qui sont importantes ou intéressantes sont transférées du registre sensoriel vers la mémoire à court terme. À ce niveau l'information est disponible pour le traitement immédiat. La taille de cette mémoire est très limitée et de plus elle

est volatile. Elle est très semblable dans ce sens à la mémoire cache dans les processeurs d'ordinateur. Elle peut retenir l'information jusqu'à 20 secondes (ou plus si l'information est rafraîchie plusieurs fois). Elle peut retenir jusqu'à 7 (plus ou moins 2) éléments. La capacité de la mémoire à court terme peut être augmentée si l'information à stocker est organisée en groupes d'informations significatifs.

- **La mémoire à long terme** : les informations qui risquent d'être utiles dans le futur sont transférées de la mémoire à court terme vers la mémoire à long terme où elles sont stockées en permanence et peuvent être retrouvées au besoin ultérieurement. La capacité de la mémoire à long terme est virtuellement illimitée. Le processus pour arriver à cette mémorisation peut prendre plusieurs aspects. Il peut être accompli par la répétition. C'est la méthode la moins efficace. Il peut être accompli par un travail cognitif supplémentaire pour trouver des liens entre les informations déjà stockées et la nouvelle information à stocker, ce qui aboutit à une meilleure rétention.
- **La notion des schémas** : les schémas sont des structures cognitives qui servent pour la représentation interne de la connaissance. Ils peuvent être combinés, étendus ou altérés au fur et à mesure de l'apprentissage.
- **L'effet de l'information significative** : l'information significative est plus facile à apprendre et à mémoriser.
- **L'effet de la position dans une série** : il est plus facile de se rappeler les éléments du début et de la fin d'une liste que les éléments du milieu de cette liste, à moins que les éléments soient très distinctifs les uns des autres.
- **L'effet de la pratique** : la pratique et la répétition améliorent la rétention et spécialement quand celles-ci sont distribuées dans le temps et dans des contextes différents.
- **L'effet de transfert** : la connaissance antérieure a un grand effet dans l'apprentissage de nouvelles informations ou de nouvelles tâches. Plus il y a de connections, plus cet apprentissage est facile.

- **L'effet de l'interférence** : la connaissance antérieure (avec l'effet de transfert) ayant une bonne incidence sur la nouvelle connaissance à apprendre (si elles sont très connectées). Cependant dans certains cas, cette connaissance peut nuire à l'apprentissage de la nouvelle connaissance. Une connaissance antérieure *semble* très liée à la nouvelle connaissance alors qu'en réalité elle ne l'est pas. Ce qui est peut être à l'origine d'erreurs de compréhension.
- **L'effet de l'organisation** : quand la connaissance à apprendre est bien structurée, son apprentissage devient plus facile.
- **L'effet de la dépendance du contexte** : quand une information est apprise dans un certain contexte, elle est plus facile à être rappelée et appliquée dans ce contexte que dans un nouveau contexte.
- **L'effet mnémonique** : des stratégies mnémoniques sont utilisées par les apprenants pour organiser en des unités significatives des informations à apprendre qui (à priori) semblent privées de sens.

Objectivité et subjectivité

L'hypothèse philosophique derrière les deux théories que nous venons d'exposer, la théorie behavioriste et la théorie cognitiviste, est qu'elles sont toutes les deux *objectivistes* : le monde observé est réel et existe d'une façon externe et indépendante de l'individu observateur [Ertmer et Newby, 1993]. Certains cognitivistes contemporains ont remis en cause cette objectivité et ont adopté une nouvelle approche subjective qui stipule que le monde n'existe qu'à travers une interprétation personnelle de ce monde et de ce fait l'apprentissage est fonction de l'individu et de son expérience. C'est l'approche constructiviste.

3.1.3 L'approche constructiviste

Le constructivisme n'est pas une approche totalement nouvelle. Comme la plupart des théories d'apprentissage, le constructivisme a des sources multiples dans les points de vues philosophiques et psychologiques du siècle dernier, et spécialement dans les travaux de Piaget, Bruner et Goodman [Perkins, 1991].

Du point de vue pédagogique, le constructivisme insiste sur le fait que l'apprentissage est un processus *actif* et *constructif*. Dans ce modèle, l'apprentissage n'a lieu que quand l'apprenant traite activement l'information présentée et en crée sa propre représentation. Ce processus de création mentale se fait nécessairement sur une base constituée des connaissances antérieures de l'apprenant qui tente de trouver des liens et des points d'ancrage pour lier toute nouvelle connaissance à la connaissance préexistante dans cette base. Du fait de l'interprétation de la nouvelle connaissance en fonction de la connaissance préexistante, les représentations mentales sont très subjectives. Et comme chaque apprenant a une expérience personnelle généralement différente des autres apprenants, et par conséquent une connaissance préexistante différente, la nouvelle connaissance va être liée différemment, et donc comprise et interprétée différemment.

Le constructivisme ne partage pas la vision du cognitivisme sur le fait que la connaissance est indépendante de l'esprit (ou du mental). Mais il ne nie pas non plus l'existence du monde réel. Seulement il suppose que ce que nous connaissons du monde provient de nos propres expériences et interactions avec ce monde. Dans cette optique, l'homme *crée* sa propre compréhension du monde à travers sa propre expérience, par opposition à *l'acquisition* pure et simple de cette connaissance.

Constructivisme et pédagogie

Dans les trente dernières années, virtuellement toutes les sciences sociales et les sciences humaines se sont éloignées du rationalisme pur et de la pensée linéaire, pour se rapprocher de l'appréciation des perspectives multiples et du raisonnement contextuel [Tarnas, 1991]. La même chose a été observée dans la conception pédagogique avec son mouvement constructiviste. Les concepteurs pédagogiques ont commencé à adopter des modèles plus flexibles et des réflexions plus polyvalentes sur l'apprentissage et l'instruction; et l'adoption de la vision constructiviste de l'apprentissage aide beaucoup dans ce sens.

[Wilson et al., 1995] définissent les concepts pédagogiques constructivistes dans les termes suivants :

- Le constructivisme est une philosophie ou une façon de penser la conception et non une approche spécifique pour la conception. Les modèles traditionnels et les stratégies pédagogiques peuvent être adaptés pour une conception constructiviste.

- Les constructivistes tentent d'éviter les recettes simples et toutes faites et ils considèrent qu'il n'y a pas de réponse simple dans la conception pédagogique.
- Les constructivistes ne sont pas des constructeurs de systèmes dans la grande tradition des théories universelles. Comme ils considèrent que la connaissance est liée aux pratiques et qu'elle est dépendante du contexte d'utilisation, leurs théories sont plus locales, partielles, et provisoires. De ce fait, très peu de ces théories sont généralisables ou utilisables d'une façon systémique.

David Merrill, bien que ne partageant pas tout le credo constructiviste, a présenté un bon modèle pour la création de systèmes éducatifs basés sur cette théorie [Merrill, 1991] :

- La connaissance est construite de l'expérience.
- L'apprentissage est une interprétation personnelle du monde.
- L'apprentissage est un processus actif de construction du sens à travers l'expérience.
- L'apprentissage est collaboratif dans le sens que la compréhension est négociée dans des perspectives multiples.
- L'apprentissage devrait être situé dans des contextes réalistes.
- L'évaluation doit être intégrée avec la tâche d'enseignement et non une activité séparée.
- La réflexion est un élément clé de l'apprentissage
- Comme la formation, l'évaluation doit être basée sur des perspectives multiples.

3.1.4 **Béhaviorisme, cognitivisme et constructivisme**

En résumé, la tendance béhavioriste considère l'apprentissage dans une vision comportementale où on ne cherche qu'à renforcer chez l'apprenant, et par le biais du conditionnement, la relation entre un stimulus externe et une réaction. L'aspect mental se résumant en une boîte noire (opaque) est complètement ignoré. Ceci est analogue à la théorie du modèle de l'atome de la fin du 19^{ième} siècle (la théorie atomique de Dalton, 1803) qui considérait le comportement de l'atome sans considérer ses constituants (Figure 3-1).

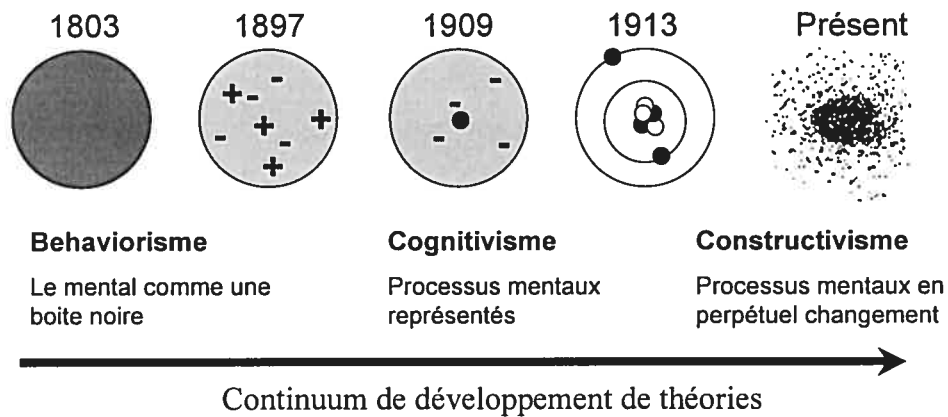


Figure 3-1 – Analogie avec le modèle de l'atome

La tendance cognitiviste considère la connaissance comme une représentation mentale symbolique et abstraite dans la tête de l'apprenant et suggère que l'apprentissage, comme le changement dans le comportement observé, ne sont que le fruit des transformations que subissent ces représentations. Pour faire l'analogie avec l'évolution de la théorie du modèle de l'atome, la tendance cognitiviste peut correspondre au modèle de l'atome du début du 20^{ème} siècle, où on a commencé à sonder l'intérieur de l'atome et à expliquer son comportement externe en fonction de sa structure interne.

La tendance constructiviste est venue éliminer la notion de représentation objective du monde que stipulaient tant la tendance behavioriste que la tendance cognitiviste. Elle a introduit la notion de subjectivité dans l'apprentissage et dans l'interprétation du monde. Un phénomène analogue c'est produit dans le monde de la physique dans la modélisation de l'atome. Au début, on considérait que les structures atomiques pouvaient être définies et représentées d'une façon très précise. Seulement, les phénomènes de la physique quantique, comme le principe d'incertitude d'Heisenberg¹ et la dualité onde-matière de l'électron par exemple, ont montré l'impossible de cette vision.

¹ Qui stipule qu'on ne peut jamais attribuer à un corpuscule à un instant donné, une position et un état de mouvement parfaitement déterminé.

| Théories d'apprentissage | Behaviorisme | Cognitivisme | Constructivisme |
|---------------------------------|---|---|--|
| Applications | Entraînement, ex. simulateur de vol. | Traitement en profondeur: exploration, organisation, synthèses. | Apprentissage collaboratif. |
| Rôle de l'instructeur | L'instructeur conçoit l'environnement d'apprentissage. | L'instructeur gère les activités de résolution de problème et les activités de recherches, spécialement en groupes. | L'instructeur surveille et dirige le processus continu de construction de la connaissance. |
| Rôle de l'apprenant | Essentiellement passif, il répond juste au stimulus | L'apprenant traite, organise, mémorise et se rappelle la connaissance apprise. | L'apprenant crée son propre apprentissage basé sur sa connaissance passée. |
| Nature d'apprentissage | -l'apprentissage est le changement dans le comportement observé -l'apprentissage est fonction des stimulus présentés -l'apprentissage est le renforcement de l'association entre le stimulus et le bon comportement | -l'apprentissage est le changement dans l'état de la connaissance et non du comportement -l'acquisition de la connaissance est une activité mentale comportant sa codification et sa structuration par l'apprenant -la connaissance antérieure de l'apprenant est prise en compte | -l'apprentissage est la création du sens par l'expérience -l'esprit filtre les informations en provenance du monde et produit sa propre réalité -l'apprenant crée une interprétation personnelle du monde en fonction de ses expériences et de ses interactions. |
| Avantage | Approprié aux habiletés motrices. | Approprié aux habiletés de résolution de problème. | Approprié aux apprentissages avec interaction et dialogue. |

Table 3-ii – Comparaison des théories behaviorisme, cognitivisme et constructivisme

3.2 Conception pédagogique et théories d'apprentissage

La conception pédagogique, que ce soit dans la période d'influence behavioriste ou dans la dernière période d'influence cognitive et constructiviste, tente généralement de se baser sur des théories d'apprentissage pour guider le processus de conception. Nous allons dans cette partie du chapitre examiner deux des théories d'apprentissage les plus influentes dans

cette discipline : la théorie des conditions d'apprentissage de Gagné et la théorie des transactions d'enseignement de Merrill.

3.2.1 La théorie des conditions d'apprentissage de Gagné

La *théorie des conditions d'apprentissage* de Gagné est la théorie d'apprentissage dominante dans le domaine de la conception pédagogique. Les premiers travaux de Robert Gagné qui touchaient à la conception pédagogique et initialement publiés dans la première édition de son livre "*The condition of learning*" [Gagne, 1965] étaient fortement enracinés dans l'approche behavioriste. Cependant, dans les éditions suivantes de son livre (Gagné, 1970, 1977, 1985), il s'est rapproché de plus en plus de l'approche cognitiviste en intégrant quelques-unes des théories de la psychologie cognitiviste, et spécialement la théorie du modèle de la cognition basé sur le traitement de l'information¹. Selon ce modèle, la cognition est un processus interne qui est sujet aux influences provenant des événements externes. Gagné conclut que l'instruction (ou l'enseignement) est l'arrangement d'événements externes qui vont activer et supporter les processus de cognition interne :

"This model posits a number of internal processes that are subject to the influence of a variety of external events. The arrangement of external events to activate and support the internal processes of learning constitutes what is called instruction" [Gagné, 1974].

"Thus it is reasonable to define instruction as being made up of events external to the learner which are designed to promote learning" [Gagné, 1974].

La théorie des conditions d'apprentissage de Gagné suggère qu'il y a plusieurs types ou natures d'apprentissage. Gagné a identifié cinq grandes catégories d'apprentissage : les informations verbales, les habiletés intellectuelles, les stratégies cognitives, les habiletés motrices et les attitudes.

- **les informations verbales** : ce type d'apprentissage est relatif à la mémorisation et au rappel de connaissances déclaratives telles les faits, les concepts, les principes et les procédures.

¹ The information-processing model of cognition

- **les habiletés intellectuelles** : ce type d'apprentissage est relié à l'acquisition de capacités intellectuelles qui peuvent être de cinq natures : la discrimination, le concept concret, le concept défini, la règle et la règle d'ordre supérieur.
 - **la discrimination** : est la capacité de distinguer les objets, les caractéristiques ou les symboles. Exemple : arriver à distinguer les notes jouées sur un instrument musical.
 - **le concept concret** : est la capacité d'identifier des classes d'objets concrets, de caractéristiques ou d'événements.
 - **le concept défini** : est la capacité de classer des idées ou des événements par leurs définitions.
 - **la règle** : est la capacité d'appliquer des règles simples dans le cadre d'une résolution de problème.
 - **la règle d'ordre supérieur** : est la capacité de combiner plusieurs règles simples pour la résolution d'un problème complexe.
- **les attitudes** : sont des états internes à l'individu qui peuvent influencer son comportement.
- **les stratégies cognitives** : ce type d'apprentissage est lié à l'acquisition de méthodes et de façons de faire, pour guider d'autres activités cognitives, tels l'apprentissage, la mémorisation, la réflexion, etc.
- **les habiletés motrices** : ce type d'apprentissage est lié à des activités physiques qui font intervenir les muscles et les mouvements coordonnés des membres.

L'importance de cette classification vient du fait que pour chacun des types d'apprentissage énoncés, Gagné a identifié les conditions particulières à respecter pour favoriser et faciliter cet apprentissage. Par exemple, pour l'apprentissage d'une stratégie cognitive, il faut donner à l'apprenant la possibilité de la pratiquer plusieurs fois et dans des contextes différents; et pour l'apprentissage d'une attitude donnée, l'apprenant doit être exposé à des situations très crédibles et qui reproduisent très fidèlement le contexte où cette attitude doit être démontrée.

En plus de cette classification, la théorie de Gagné a identifié neuf événements pédagogiques ordonnés sur lesquels peut s'organiser la conception pédagogique (Figure 3-2). Ces événements pédagogiques sont :

1. **Attirer l'attention** : Aucun apprentissage n'est possible sans l'attention de l'apprenant. Attirer l'attention de l'apprenant et la maintenir est essentiel pour la réussite d'un apprentissage. Les techniques et les tactiques sont très variées. Ils peuvent prendre la forme de questions de réflexion plus ou moins sophistiquées, comme une présentation multimédia pour situer l'apprenant et piquer sa curiosité.
2. **Annoncer les objectifs de la formation** : chaque session d'apprentissage doit indiquer à l'apprenant les objectifs de formation de la session. L'énoncé de ces objectifs permet à l'apprenant d'anticiper le contenu de la session, ce qui facilite l'acquisition et l'intégration de la connaissance nouvelle qui va être présentée.
3. **Stimuler le rappel des préalables** : l'association des nouvelles informations avec la connaissance et l'expérience antérieure de l'apprenant peut faciliter l'apprentissage. Le rappel de la connaissance antérieure peut être explicite par l'insertion des éléments à rappeler dans le contenu actuel, ou implicite par l'évocation de questions se rapportant à cette connaissance antérieure.
4. **Présenter le nouveau contenu** : c'est dans cet événement pédagogique que le matériel didactique est effectivement présenté à l'apprenant. Le contenu doit être subdivisé et organisé d'une manière significative.
5. **Guider l'apprentissage** : pour aider l'apprenant à mieux comprendre le matériel didactique présenté, des conseils et des directives de soutien doivent être fournis tout au long de la présentation. Ces stratégies des conseils incluent la présentation d'exemple, de contre exemple, d'étude de cas, des astuces mnémoniques, et des analogies.
6. **Provoquer la performance** : cet événement pédagogique est concerné par la pratique. L'apprenant est invité à mettre en pratique la nouvelle connaissance ou capacité acquise. C'est une façon de confirmer la compréhension correcte de l'apprenant, et c'est aussi un moyen de favoriser la mémorisation.

7. **Fournir un feedback sur la performance de l'apprenant** : la performance de l'apprenant dans la phase de pratique doit être suivie et contrôlée. Il est essentiel pour l'apprenant de recevoir des feedbacks spécifiques et immédiats sur son travail. Cette activité ne doit pas faire partie de l'évaluation de l'apprenant mais doit servir uniquement à des fins de compréhension et de pratique.
8. **Évaluer la performance de l'apprenant** : au terme d'un module pédagogique ou d'une session, l'apprenant doit avoir l'opportunité de passer un post-test ou examen. Cet examen doit être fait sans aucune assistance. Le résultat de cette évaluation est à la base de la décision à prendre sur le prochain module à présenter.
9. **Favoriser la rétention et le transfert** : un concept peut être bien appris à un certain moment mais oublié ultérieurement. Pour faciliter la rétention et le transfert des concepts appris, des répétitions espacées dans le temps, de l'utilisation de ce concept, doivent être prévues dans la planification pédagogique.

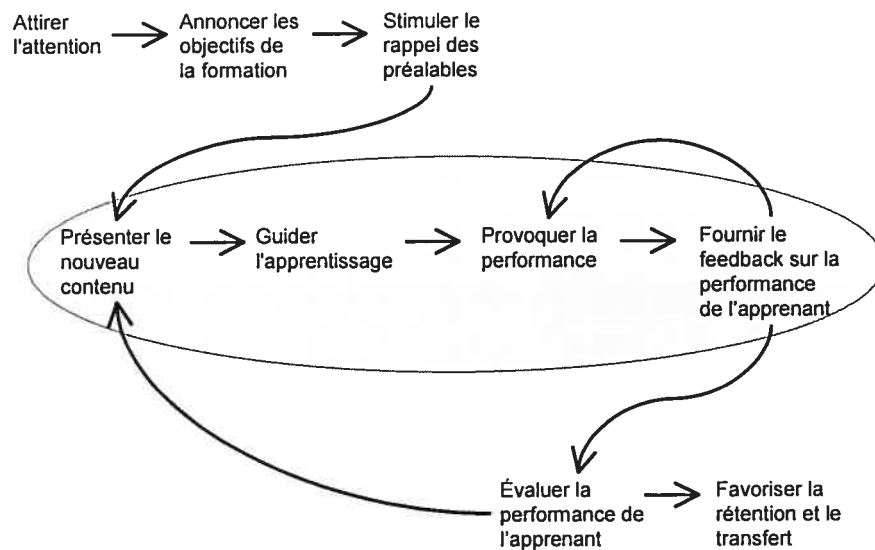


Figure 3-2 – Circuit des neuf événements pédagogiques de Gagné

3.2.2 La théorie des transactions d'enseignement de Merrill

La théorie des transactions d'enseignement de Merrill [Merrill, 2000], ou la ITT (pour *Instructional Transaction Theory*) est une amélioration d'une autre théorie très connue de Merrill, la *Component Display Theory*, ou la CDT [Merrill, 1983]. La théorie des transactions d'enseignement est basée principalement sur les mêmes bases que la théorie des conditions d'apprentissage de Gagné. Les deux théories s'accordent sur le fait qu'il y a différents types d'apprentissage et que chacun de ces types d'apprentissage nécessite différentes stratégies d'enseignement et différentes manières d'évaluation. Merrill juge que les neuf événements pédagogiques proposés par Gagné sont nécessaires dans une approche pédagogique. Cependant, la théorie des transactions d'enseignement de Merrill donne beaucoup d'importance à des aspects pratiques comme la représentation de connaissances par exemple, ce qui la rend plus apte à l'automatisation dans les systèmes tutoriels informatisés.

Le premier aspect important de la théorie des transactions d'enseignement de Merrill est sa classification explicite de la connaissance du domaine à enseigner. Merrill suggère quatre types de connaissance pour cette classification : les faits, les concepts, les processus, les procédures et principes.

- **Le fait** : est une information sur un événement ou sur les relations qui existent entre un certain nombre d'objets ou d'entités à un moment donné.
- **Le concept** : est un ensemble d'objets qui ont des caractéristiques communes et un nom commun.
- **La procédure** : est une séquence d'étapes à suivre conçue pour résoudre un problème.
- **Le principe** : est une affirmation sur les relations de causalité ou de corrélation entre phénomènes. Exemple d'un principe en photographie "*l'éclairage de côté crée de l'ombre sur le visage.*"

Merrill a classifié aussi les tâches que peut effectuer un apprenant pour acquérir cette connaissance. Il a prévu trois catégories de tâches : *se rappeler, utiliser/appliquer, trouver/généraliser*. La combinaison des deux classifications donne la matrice des tâches montrée dans la Table 3-iii.

| | Faits | Concepts | Processus | Procédures | Principes |
|----------------------------|---------------------|--|---|---|---|
| Se rappeler | Se rappeler un fait | Se rappeler une définition d'un concept | Se rappeler les étapes d'un processus | Se rappeler les étapes d'une procédure | Se rappeler les lignes directives |
| Utiliser/Appliquer | | Classifier un nouvel exemple | Résoudre un problème, faire une inférence | Exécuter une procédure | Résoudre un problème, faire une inférence |
| Trouver/Généraliser | | Définir une classe ou un groupe d'objets ou d'événements | Développer un processus | Dériver, créer une procédure ou une technique pour atteindre un but donné | Découvrir les relations de cause-à-effet |

Table 3-iii – Matrice des tâches

Pour prendre en compte, et d'une façon explicite, les aspects dynamiques et causaux qui sont très souvent présents dans les contextes d'enseignement, Merrill introduit une nouvelle perspective dans la représentation de la connaissance liée spécialement à la causalité et au dynamisme. Il suggère quatre types d'*objets de connaissance* :

- **les entités**, sont les choses et les objets du domaine. Une entité peut être composée d'autres entités qui en sont les parties.
- **les actions**, sont les procédures que peut effectuer un apprenant avec ou sur une entité ou ses parties.
- **les processus**, forment la suite d'événements qui peuvent arriver généralement à la suite d'une action. Un processus a une ou plusieurs conditions pour son déclenchement.. Si ces conditions sont vérifiées alors le processus est exécuté. Les conditions sont exprimées en fonction de valeurs de propriétés. L'exécution d'un processus a pour conséquence le changement dans les valeurs de propriétés d'un certain nombre d'entités. Ces changements peuvent affecter les pré-conditions d'autres processus qui peuvent être déclenchés à leur tour.

- **les propriétés**, sont des descriptions servant à décrire les entités, les actions et les processus. À une propriété on peut associer des valeurs qualitatives ou quantitatives. À chaque valeur d'une propriété peut être associé un portrait de cette valeur qui montre comment l'apprenant va percevoir l'état de la propriété exprimé par cette valeur. Le portrait peut être symbolique, visuel (image, animation, etc.) ou même olfactif ou cinétique.

Bien que ce modèle soit simple et comporte peu d'éléments, il est relativement efficace sur le plan pédagogique car il permet, tant au concepteur de cours qu'à l'apprenant, de se focaliser sur les aspects dynamiques et causaux et de bien les comprendre (Figure 3-3).

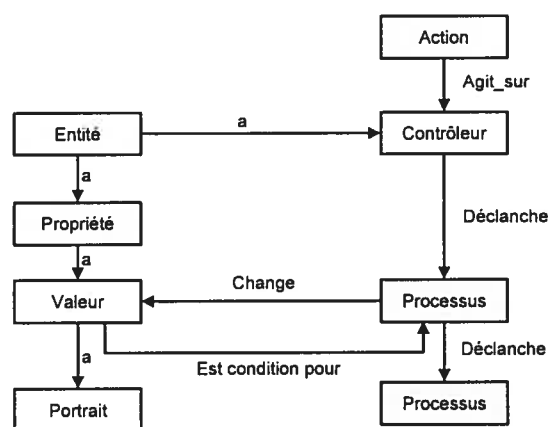


Figure 3-3 – Structure de réseau causal basé sur le modèle de Merrill

L'autre aspect important de la théorie des transactions d'enseignement de Merrill est celui de la présentation du contenu ou de l'expérience didactique. Merrill suggère que toute présentation didactique consiste dans son essence en une suite discrète d'affichages. Il a nommé ces affichages : *les formes de présentation primaires*.

Il est clair que par rapport à la théorie de Gagné, la théorie des transactions d'enseignement de Merrill s'occupe beaucoup plus du niveau de détail de l'enseignement et tient compte des aspects de la représentation des connaissances et des aspects de l'interaction avec l'apprenant. Cependant, les deux théories, comme les autres théories d'apprentissage, s'intéressent aux conditions qui forment l'environnement d'apprentissage sans s'intéresser aux mécanismes d'apprentissage spécifiquement. Dans la prochaine section de ce chapitre, nous explorons ces mécanismes et leurs relations avec la représentation de la connaissance.

3.3 Les mécanismes d'apprentissage

Vosniadou [Vosniadou, 1996] considère l'apprentissage comme un changement dans la façon avec laquelle l'apprenant (plus précisément son cerveau) traite l'information. Ce changement dans le comportement humain peut s'opérer à deux niveaux : au niveau des stratégies de traitement de l'information utilisées dans les processus cognitifs, et au niveau de la représentation de la connaissance qui supporte ces processus cognitifs.

3.3.1 Apprentissage et stratégies de traitement de l'information

Le processus d'apprentissage implique l'acquisition de nouvelles procédures pour le traitement de l'information et la révision des procédures déjà acquises. Apprendre à résoudre un problème par exemple, nécessite l'acquisition de stratégies et de procédures qui rendent possible la mise au point, la planification et l'exécution d'une solution. Les facultés intellectuelles, telle la mémorisation et la compréhension font elles aussi appel à des stratégies et des procédures. La mémorisation par exemple fait appel à des stratégies comme la répétition, la récitation et l'utilisation d'images visuelles, alors que la compréhension fait appel à des stratégies métacognitives comme la planification, le questionnement, la prédiction, etc. Mais quel que soient le domaine et la faculté concernés par un apprentissage, certains mécanismes d'apprentissage restent applicables. Parmi ces mécanismes on retrouve :

- **Le renforcement (*Strength accumulation*)** : les procédures fructueuses dans la résolution d'un problème sont soumises à des mécanismes de généralisation et de sélection ayant pour but de distinguer les caractéristiques du problème qui peuvent prédire la réussite ou l'échec d'une certaine procédure. Par ce moyen, les stratégies de résolution qui sont fructueuses se trouvent renforcées alors que celles infructueuses se trouvent amoindries.
- **Le réglage et le raffinage (*Tuning*)** : les procédures sont généralement réglées et raffinées pour être mieux appliquées dans un domaine. Ceci passe par les mécanismes de généralisation et de sélection qui font en sorte que l'application d'une procédure soit plus ou moins restreinte, dépendamment de l'historique des réussites et des échecs dans l'application de cette procédure.

- **La composition** : c'est un mécanisme d'apprentissage qui combine un ensemble de procédures en une seule procédure. Cette opération fait accroître la vitesse d'apprentissage par la création de nouveaux opérateurs qui incarnent, dans un seul ensemble consolidé, la séquence d'étapes nécessaires pour la résolution d'un problème donné.
- **La procéduralisation (*Proceduralization*)** : le mécanisme de procéduralisation remplace graduellement l'information déclarative par des procédures. L'information déclarative comme les faits et la description des règles par exemple sont plus faciles à apprendre car ils ne font intervenir principalement que la mémorisation. À l'opposé, l'apprentissage des procédures est plus complexe et nécessite généralement plus de temps, car en plus de la mémorisation des différentes étapes il faut mémoriser les conditions d'utilisation et les mécanismes d'application. Le mécanisme de procéduralisation permet de convertir l'information verbale qui décrit l'application d'une ou d'un ensemble de règles, en une procédure (compacte) que l'apprenant peut reconnaître et appliquer plus rapidement. Une fois cette procédure formée, elle peut être mise à jour par les mécanismes de renforcement et de raffinement décrits ci haut.
- **L'examen (*Inspection*)** : ce mécanisme permet l'apprentissage par l'examen et l'observation. De nouvelles règles peuvent être apprises par l'examen des solutions passées et l'observation de régularités présentes dans ces solutions.
- **L'analogie (*Analogy*)** : le mécanisme d'apprentissage par analogie permet le transfert et l'adaptation d'une solution d'un problème donné et la proposer comme solution pour un problème similaire dans le même domaine ou dans un domaine différent.

Ces mécanismes d'apprentissage sont génériques et indépendants d'un domaine d'application particulier. Les stratégies et les procédures qu'ils créent ou modifient touchent essentiellement le processus de traitement de l'information sans tenir compte de comment ce traitement affecte et réorganise cette information. C'est que nous allons montrer dans les paragraphes suivants.

3.3.2 Apprentissage et représentation de la connaissance

La psychologie cognitive s'est longtemps intéressée à la façon dont la connaissance est organisée au niveau de la mémoire à long terme chez l'humain. Différentes

conceptualisations et théories ont été proposées pour expliquer les différents phénomènes qui caractérisent l'activité de l'intelligence humaine. Parmi ces principes on retrouve les notions de *concepts*, de *structures conceptuelles*, de *schémas*, de *frames*, etc. Chacun de ces principes, du point de vue de l'apprentissage, peut faire appel à des mécanismes différents.

3.3.2.1 Les concepts et les structures conceptuelles

Il est généralement supposé que les personnes se représentent le monde à travers une série de concepts qui sont organisés dans une large structure conceptuelle. Dans les premières théories proposées pour expliquer la nature du *concept*, comme dans la vision classique (*the classical view*), le concept est décrit en terme d'un ensemble d'attributs (nécessaires et suffisants) qui le définissent en spécifiant laquelle de ses instances appartient à une catégorie donnée et laquelle n'y appartient pas. Avec la même vision Smith et al. [Smith et al., 1974] supposaient que les concepts sont représentés mentalement par deux types d'attributs, les traits définitoires et les traits caractéristiques:

- Les traits définitoires : les traits nécessaires et suffisants pour déterminer l'appartenance catégorielle d'un concept.
- Les traits caractéristiques : les traits qui appartiennent aux exemplaires les plus typiques du concept mais qui n'appartiennent pas nécessairement à tous les exemplaires (ou les individus) de ce concept.

[Denhière et Baudet, 1992] utilisent le terme '*trait sémantique*' pour définir ce qui caractérise un concept et considèrent que l'apprentissage d'un concept consiste à identifier et à mémoriser ces traits sémantiques :

“ ... les concepts sont représentés par un ensemble de traits sémantiques correspondants à un sous-ensemble du savoir que possèdent les individus sur une catégorie ou sur un individu. ... Les stimulus auxquels sont confrontés les être humains se différencient par des caractéristiques multiples et changeantes. Une des fonctions de l'apprentissage consiste à extraire de cet environnement mouvant des invariants – des propriétés – et à en construire une trace mémorielle plus au moins permanente : les traits sémantiques. ... la fonction essentielle d'un trait sémantique est de constituer un support de discrimination”.

Les traits qui définissent un concept indépendamment du nom qu'on leur donne – définitoires, caractéristiques ou sémantiques – font généralement référence à d'autres

concepts. Par exemple pour définir un triangle on utilise le trait caractéristique “avoir trois côtés”. Ce trait caractéristique fait référence à son tour au concept “trois” et au concept “côté”. C’est pour cette raison que *la théorie associationniste* [Luger et Stubblefield, 1998], suggère que ce qui est important dans la définition d’un concept, c’est le réseau d’associations qu’il a avec les autres concepts.

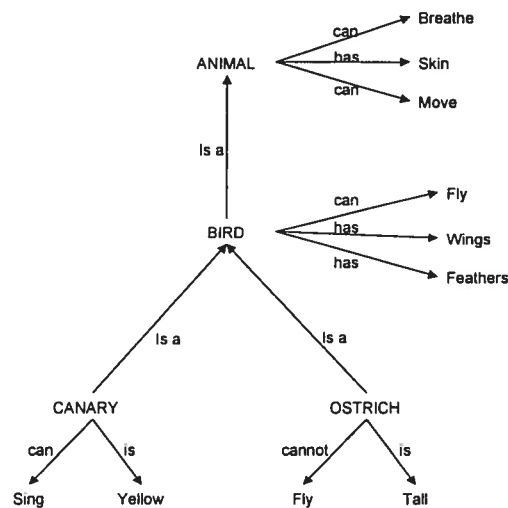


Figure 3-4 – Exemple de réseau sémantique [Collins et Quillian, 1969]

Les travaux de [Collins et Quillian, 1969] ont démontré que la création de ces associations aboutit à la création d’une structure hiérarchique et que cette structure conceptuelle peut être modélisée par un réseau sémantique comme dans l’exemple de la Figure 3-4. Les psycho-cogniticiens pensent que plusieurs changements peuvent affecter cette structure à la suite d’un apprentissage. Les changements les plus simples peuvent consister en des ajouts ou des suppressions d’attributs pour un concept donné. Les changements les plus complexes peuvent consister en des ajouts ou des suppressions de relations entre concepts ou à des changements dans l’emplacement des concepts et dans l’organisation de la structure conceptuelle.

3.3.2.2 Les schémas

Une autre théorie proposée pour expliquer l’organisation des concepts dans la mémoire est la théorie des frames ou des schémas [Minsky, 1975]. Les schémas sont supposés encoder la connaissance générique qui est applicable dans des contextes et des situations différentes. Ils sont structurés par l’expérience humaine et sont rappelés et adaptés au besoin. Les schémas

sont supposés représenter une connaissance relativement abstraite qui caractérise la connaissance initiale (*the background knowledge*) que possèdent les gens sur des situations stéréotypées. [Ohlsson, 1996] suggère que les schémas sont des modèles (ou patterns) avec une certaine structure interne et qui jouent un rôle important dans la pensée et le raisonnement et postule que l'acte de raisonner revient à aligner des patterns ("*To think is to align patterns*").

L'intérêt des schémas pour l'apprentissage est qu'ils sont très pratiques pour capturer le processus vertical de décomposition et d'interprétation de l'information présentée aux apprenants et pour expliquer les erreurs survenant souvent dans de telles situations.

Les types d'apprentissages qu'on peut observer basés sur le modèle des schémas ont été décrits par [Rumelhart et Norman, 1978] qui en ont distingué trois : l'apprentissage par accumulation, l'apprentissage par réglage et raffinement, et l'apprentissage par restructuration.

Dans l'apprentissage par accumulation, les schémas sont graduellement augmentés par l'ajout de nouveaux faits. L'apprentissage par réglage et raffinement a lieu lors de la pratique répétée de l'utilisation d'un schéma et réfère aux différentes améliorations qui peuvent être apportées à ce schéma comme la généralisation ou la restriction de son utilisation et la détermination de ses valeurs par défaut (valeurs initiales). Finalement, l'apprentissage par restructuration réfère aux différents changements dans la structure d'un schéma pour l'adapter à une nouvelle situation, ou à la fusion de plusieurs schémas pour la création d'un nouveau schéma mieux adapté à un nouveau contexte. Ohlsson offre une vision similaire et considère que l'apprentissage réalisé par un apprenant a pour effet de faire évoluer le schéma mental de cet apprenant d'un état simple (et potentiellement invalide) qui caractérise généralement le niveau d'un novice, à un état de plus en plus complexe et de plus en plus riche, qui caractérise généralement le niveau d'un expert (Figure 3-5). Ohlsson ne considère pas le schéma initial de l'apprenant, s'il est invalide, comme faux ou erroné, mais plutôt comme une interprétation différente de celle de l'expert. Ceci va dans le même sens que la théorie constructiviste de l'enseignement.

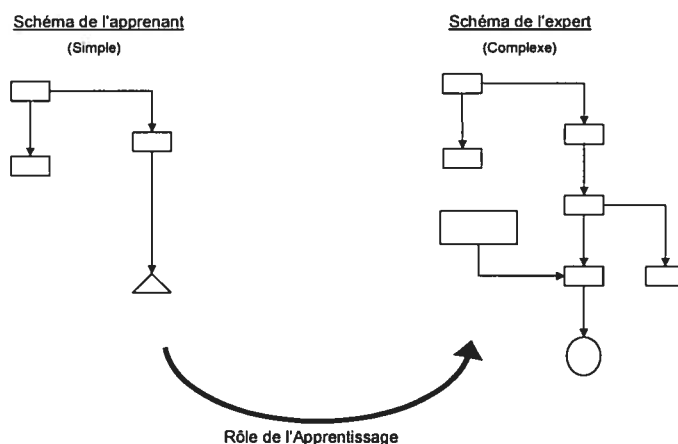


Figure 3-5 – Rôle de l'apprentissage dans la théorie des schémas

3.3.2.3 Les théories et les modèles mentaux

Sur le plan cognitif, les concepts, les structures conceptuelles et les schémas sont souvent intégrés dans des *théories* ou des *modèles mentaux* pour expliquer et raisonner sur le monde.

Le terme théorie réfère à l'image du monde que se fait l'apprenant et à une notion moins forte que dans la définition de la théorie scientifique conventionnelle qui implique des notions beaucoup plus rigoureuses comme la systémativité, l'abstraction et l'objectivité. La notion de théorie a été introduite pour faire la distinction entre :

- un ensemble cohérent de connaissances qui expliquent le monde et les relations de causalité observées,
- et un ensemble de connaissance constitué uniquement de faits.

Une théorie peut être vue comme un schéma d'un type spécial, qui a pour but de fournir un cadre où les phénomènes peuvent être compris et interprétés. Et comme les schémas, les théories sont modelées par l'apprentissage et ces mécanismes d'accumulation, de réglage et raffinement, et de restructuration.

Les modèles mentaux sont des représentations conceptuelles et opérationnelles que les humains construisent quand ils interagissent avec un système complexe [Jonassen, 1995]. Ce sont des représentations analogues qui préservent la structure de ce qu'ils représentent et qui sont utilisés comme base pour construire des explications et des déductions qui ne sont pas

directement disponibles dans la connaissance déjà mémorisée. Pour cette raison on dit qu'ils sont génératifs et dynamiques.

Il est supposé que la plupart des modèles mentaux sont construits à la volée et au besoin pour faire face à une situation particulière (résoudre un problème ou répondre à une question non familière), cependant il est possible que certains modèles, fréquemment utilisés, soient stockés dans la mémoire à long terme. Le but de l'apprentissage et de la pratique des exercices, peut parfois être la création de modèles mentaux valides nécessaires pour répondre à une situation et leurs transferts avec la pratique et la répétition, de la mémoire à court terme (ou la mémoire de travail) où ils sont initialement (re-)construits, à la mémoire à long terme où ils sont stockés en permanence et d'où ils vont être *instanciés* plus rapidement (Figure 3-6).

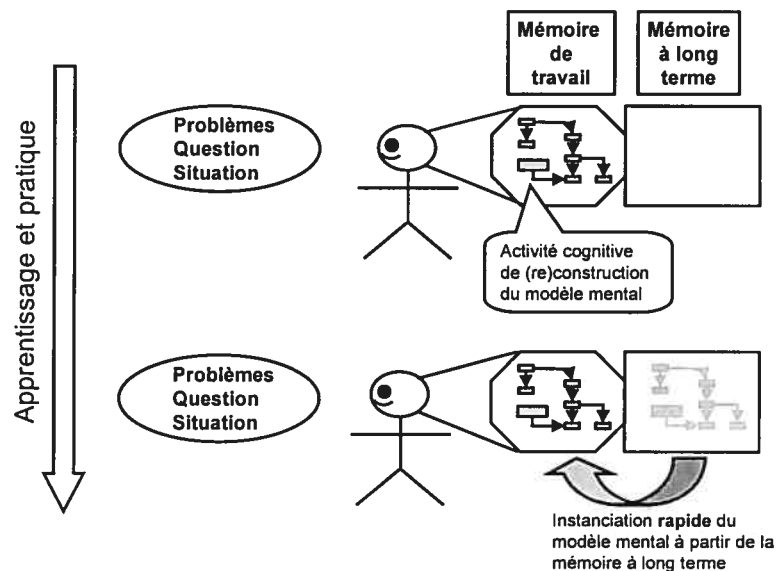


Figure 3-6 – Apprentissage et modèles mentaux

Les modèles mentaux ont plusieurs fonctions dans le système cognitif humain. [Vosniadou, 2002] a distingué trois fonctions importantes : a) comme aide à la construction des explications, b) comme médiateur dans l'interprétation de nouvelles informations, et c) comme outils qui permettent l'expérimentation mentale des théories et leur révision. La Figure 3-7 donne quelques exemples de modèles mentaux faits par des élèves en école primaire qui montrent leur compréhension de la forme de la terre [Vosniadou et Brewer, 1992].

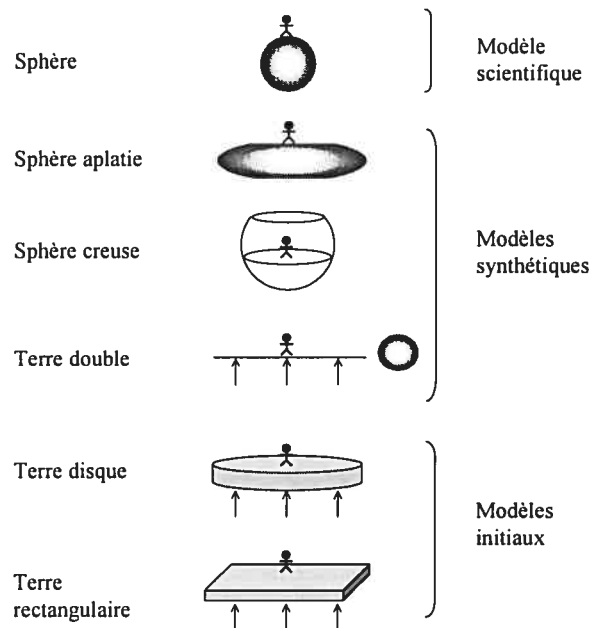


Figure 3-7 – Modèle mental pour la terre [Vosniadou et Brewer, 1992]

Dans un processus de raisonnement humain, des modèles mentaux sont construits surtout dans les situations où les conclusions ne peuvent pas être directement trouvées ou déduites de l'information mémorisée. Les modèles mentaux sont alors d'une aide précieuse car ils transforment la connaissance implicite (présente mais non évidente) en une connaissance explicite directement exploitable et utilisable, particulièrement dans la génération d'explications.

Les modèles mentaux ont aussi un rôle important dans l'interprétation de nouvelles informations qui surviennent. Une fois un modèle mental construit (ou retrouvé de la mémoire à long terme), s'il est retenu comme contexte adéquat pour l'information nouvelle, il devient une sorte de structure-contrainte à laquelle la nouvelle information doit s'associer en cherchant à y trouver sa place : c'est une interprétation de la nouvelle connaissance en fonction du modèle mental instancié. Les erreurs de compréhension et les contradictions chez les apprenants peuvent généralement être expliquées comme une tentative d'intégrer une nouvelle information dans un modèle mental existant qui n'est pas approprié pour la circonstance ou qui n'est pas mis à jours et adapté pour pouvoir accepter la nouvelle information.

3.4 Connaissance du domaine et apprentissage de haut niveau.

Les concepteurs pédagogiques ont longtemps reconnu l'importance de l'analyse du sujet à enseigner à des fins de faciliter l'apprentissage par *la sélection, l'organisation et le séquençement* appropriés de la connaissance du domaine:

“...it is our hypothesis that the greatest impact on learning results from the representation and organization of the knowledge to be learned.”
[Merrill, 2002].

Cette analyse et organisation de la connaissance du domaine prennent encore plus d'importance dans le cadre de l'automatisation de la formation au niveau des systèmes tutoriels intelligents.

3.4.1 Nature de la connaissance du domaine

Pour expliquer la nature de la connaissance du domaine à enseigner, plusieurs suggestions ont été avancées afin de fournir des définitions et des classifications de son contenu. Selon Gagné, la connaissance du domaine peut être formée des composantes suivantes [Gagné, 1985] :

- les faits,
- les concepts,
- les règles (les procédures simples) et
- les règles d'ordre supérieures (les procédures compliquées).

Reigeluth, Merrill et Bunderson [Reigeluth et al., 1978] identifient les composantes de la connaissance du domaine comme étant:

- les faits,
- les concepts,
- les étapes (d'une procédure),
- les principes,
- les listes : pour représenter une collection ou un ensemble d'éléments,

- la taxonomie des parties : une structure hiérarchique pour représenter les éléments et leurs composants ou leurs parties,
- la taxonomie ontologique : une structure hiérarchique pour représenter les types d'objets et les catégories par les relations de genre *est-um*,
- les pré-requis procéduraux : des structures qui organisent les étapes d'une procédure avec l'ordre à respecter entre les étapes,
- les décisions procédurales : des structures qui identifient les alternatives et les facteurs que doit prendre l'apprenant en compte pour prendre une décision du choix entre les alternatives,
- Les cause-à-effet : des structures qui identifient les relations entre les causes et leurs effets.

Gagné et Merrill ont étendu cette structuration de la connaissance [Twitchell, 1990] et ils ont proposé les structures suivantes :

- la liste,
- les taxonomies : type, partie, propriété, fonction,
- les algorithmes : chemin, décision, et
- les réseaux causaux : chaînes d'événements, chaînes de causes.

Ohlsson dans sa définition pour la connaissance du domaine [Ohlsson, 1996], fait la distinction entre la connaissance pratique (ou la connaissance procédurale) et la connaissance déclarative :

- La connaissance pratique est la connaissance qui permet d'agir et montre comment effectuer une tâche ou atteindre un but désigné. Cette connaissance peut être reliée à des habiletés sensori-motrice (e.g. "*Conduire une voiture.*") ou à des habiletés cognitive (e.g. "*Jouer aux échecs.*").
- La connaissance déclarative est une connaissance qui permet de se représenter le monde. Elle peut concerner des faits concrets (e.g. "*La lune n'a pas d'atmosphère.*"), ou une connaissance abstraite comme les principes (e.g. "*La loi du mouvement de Newton.*") ou les idées (e.g. "*L'idée du champ gravitationnel.*").

Les dimensions de la connaissance du domaine.

C'est qui est important à retenir des ces classifications de la connaissance du domaine que nous venons d'exposer, c'est la nature de la connaissance qu'elles cherchent à capter. Nous avons combiné les différentes visions proposées et nous avons abouti à six dimensions pour la classification et l'organisation de la connaissance du domaine (Figure 3-8). Ces dimensions sont :

- **La dimension factuelle** : relative aux faits.
- **La dimension conceptuelle** : relative aux concepts, aux taxonomies ontologiques, aux taxonomies des types et aux taxonomies des propriétés.
- **La dimension structurelle** : relative aux collections, aux listes, aux taxonomies des parties, etc.
- **La dimension procédurale** : relative aux règles, aux étapes et aux algorithmes.
- **La dimension causale** : relative aux relations de cause à effet et aux réseaux causaux.
- **La dimension abstraite** : relatives aux principes et aux idées.

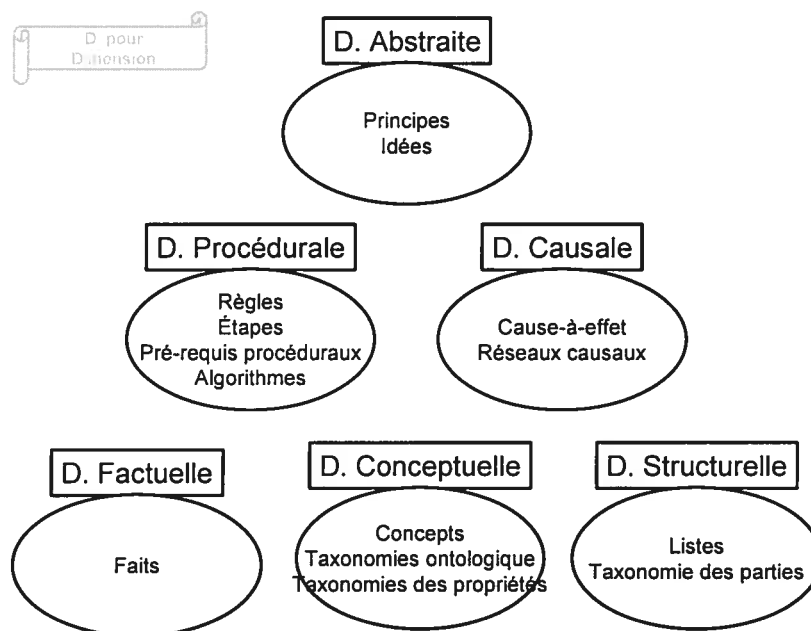


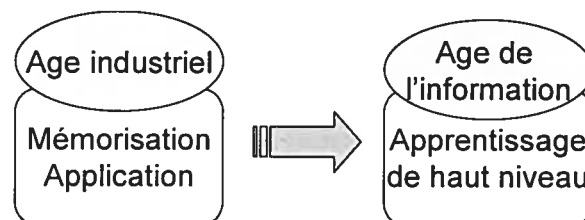
Figure 3-8 – Pyramide de la connaissance du domaine

Dans cette définition des dimensions de la connaissance, nous faisons abstraction des différences qui existent entre la modélisation de la connaissance chez l'humain et la modélisation de la connaissance aux niveaux des STIs. Ces derniers par exemple, en tant que systèmes informatiques, ont besoin d'une connaissance formelle, concise et très détaillée, alors que l'être humain est capable d'utiliser une connaissance informelle, redondante et grossière.

3.4.2 Apprentissage de haut niveau

La nature de l'apprentissage recherché, tant dans l'enseignement conventionnel que dans l'enseignement basé sur les STIs, est resté longtemps limité exclusivement aux aspects de la mémorisation pour la connaissance déclarative, et aux aspects de l'application pour la connaissance procédurale [Reigeluth et Moore, 1999] [Knight, 2002] [Ohlsson, 1996]. Cette vision de l'éducation est due essentiellement aux lois économiques qui veulent que l'éducation en général, et au niveau supérieur en particulier, soit plus concentrée sur une connaissance applicable dans le monde réel et donne moins d'importance à la connaissance de réflexion [Symes et McIntyre, 2000]. Beaucoup d'auteurs ont commencé à critiquer cet état de fait et à réclamer la nécessité de prendre en compte un troisième aspect de la nature d'apprentissage : l'apprentissage de haut niveau ou l'apprentissage complexe.

“In the industrial-age paradigm of instructional-design theory, work focused almost exclusively on the cognitive domain, and within that domain, it focused almost exclusively on remember-level and application-level learning (memorization and procedural-skill development). While information technologies and information-age roles make those levels of learning less important now, they still have important place. But higher levels of learning are becoming relatively much more important for the vast majority of learners. And there is a greater need for methods of instruction that allow for much greater customization of the learning experience and much greater utilization of information technology, fellow learners, and other resources for learning.” [Reigeluth et Moore, 1999]



[Ohlsson, 1996] soutient que l'apprentissage de haut niveau est basé essentiellement sur la connaissance déclarative abstraite, telle les idées et les principes, et affirme que la compréhension profonde d'un sujet passe nécessairement par ce type de connaissance. Ohlsson ajoute que la majorité de la connaissance humaine est de nature déclarative et abstraite, plutôt que de nature pratique (ou procédurale).

[Claxton, 1998] suggère que ce type d'apprentissage a tendance à être lent, et que généralement il dépasse le cadre d'une session d'apprentissage. Ceci implique que la conception pédagogique doit prendre en compte ce facteur et distribuer l'acquisition de telles connaissances sur de longues périodes. Ceci implique aussi qu'il faut penser la conception dans sa globalité et non modules par modules ou unités par unités, qui sont relativement trop courts pour permettre l'apprentissage de haut niveau.

[Biggs, 1999] pense qu'un bon apprentissage, celui qui favorise l'apprentissage de haut niveau dans le contexte de l'éducation supérieure, exige la présence d'activités d'enseignement et d'apprentissage qui encourage la compréhension profonde. Les activités pédagogiques de bas niveau, comme la mémorisation, l'identification, l'énumération et l'application de règles et de procédures, sont nécessaires mais elles ne peuvent pas suffire à elles seules. Les apprenants ont besoin d'être impliqués dans des activités intellectuelles de haut niveau comme l'argumentation, l'explication, la réflexion, l'identification de relations interconnaissance et le transfert et l'adaptation des connaissances et procédures entre cas analogues (Figure 3-9).

[Ohlsson, 1996] qui s'est inspiré de la théorie des jeux et des formes épistémiques¹ de [Collins et Ferguson, 1993] a proposé une liste d'activités, qu'il a nommées tâches épistémiques, et qui peuvent favoriser l'apprentissage de haut niveau (Table 3-iv) : *Décrire, Expliquer, Prédire, Argumenter, Critiquer, Clarifier, Définir*.

¹ The *epistemic forms and epistemic games*

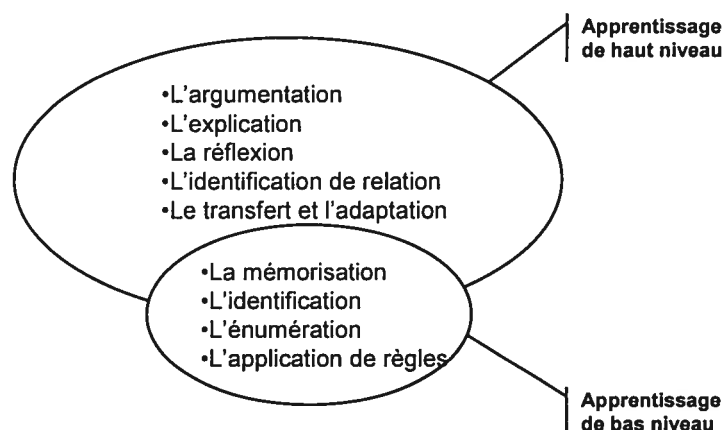


Figure 3-9 – Apprentissage de bas niveau vs. de haut niveau

| Tâche épistémique | Description |
|-------------------------|--|
| Describing | To fashion a discourse referring to an object or event such that a person who partakes of that discourse acquires an accurate conception of that object or event |
| Explaining | E.g. in relation to some event, to fashion a discourse such that a person who partakes of that discourse understands why that event happened |
| Predicting | To fashion a discourse such that a person who partakes of that discourse becomes convinced that such and such an event will happen |
| Arguing | To state reasons for (or against) a particular position on some issue thereby increasing (or decreasing) the recipient's confidence that the position is right. Arguments are about what to believe and what to do |
| Critiquing (evaluating) | To critique a cultural product is to fashion a discourse such that a person who partakes of that discourse becomes aware of the good and bad points of that product |
| Explicating | To explicate a concept is to fashion a discourse such that a person who partakes of that discourse acquires a clearer understanding of its meaning |
| Defining | To define a term is to propose a usage for that term. |

Table 3-iv – Définition des tâches épistémiques selon Ohlsson

3.5 Conclusion

Nous avons exposé dans ce chapitre les tendances majeures en conception et en psychologie pédagogique. Elles insistent sur le fait que l'apprentissage doit être de nature constructiviste avec la pleine participation de l'apprenant dans le processus de la construction de la connaissance. Nous pensons que pour mettre en pratique cette vision constructiviste dans le cadres des STIs, il faut rendre le processus de construction de la connaissance explicite et comme faisant partie naturelle de l'expérience d'enseignement.

Pour ceci, il faut prendre en compte la nature de la connaissance qui compose le domaine. Cette connaissance doit être représentée pour supporter le processus de construction. Seulement, nous pensons que le niveau et le détail de cette connaissance doit être de nature à supporter l'apprentissage de haut niveau, tels les idées et les principes, et ainsi permettre d'engager l'apprenant dans des activités intellectuelles, telles que les tâches épistémiques, qui vont au-delà du simple rappel de la connaissance ou de l'application d'une procédure.

[Ohlsson, 1996] suggère que la connaissance déclarative abstraite, celle qui permet l'apprentissage de haut niveau, ne se manifeste pas dans les pratiques procédurales, mais essentiellement dans le discours. Cette vision est confirmée par l'observation que l'on peut faire au niveau de l'enseignement supérieur, où malgré toutes les avancées au niveau de la technologie, les étudiants sont beaucoup plus à l'aise avec un livre qu'avec un autre moyen électronique, quand vient le temps d'apprendre et réfléchir profondément dans un domaine. Ceci est dû essentiellement à cette composante discours dans le livre éducatif et à laquelle nous sommes habitués [Weller, 2000].

Dans le prochain chapitre nous proposons une nouvelle approche pour la modélisation et la représentation de la connaissance, tant du domaine que pédagogique, destinée essentiellement à tirer profit du livre éducatif et de son discours, pour promouvoir l'apprentissage de haut niveau.

CHAPITRE 4

LE MODÈLE *LEKC*: POUR LE DISCOURS, L'ÉPISTÉMIQUE ET LA DIDACTIQUE

4.1 Introduction

L'importance et la portée de chaque nouvelle invention ou innovation, selon Winograd et Flores [Winograd et Flores, 1987], sont fortement tributaires de la façon dont elle s'insère dans le réseau existant des pratiques, des moyens, des institutions et des conventions.

La lecture, l'annotation et les ressources textuelles, comme pratiques et moyens, sont très largement utilisées pour la découverte et l'acquisition de connaissances à tous les niveaux du système éducatif et particulièrement dans l'éducation supérieure.

L'approche de modélisation de cours que nous proposons dans ce chapitre, nommée *LEKC*, pour *Learning by Explicit Knowledge Construction Framework* [Rouane, Frasson et Kaltenbach, 2002-a] [Rouane, Frasson et Kaltenbach, 2003-b], est conçue pour s'insérer dans les pratiques courantes de lecture en contexte d'apprentissage pour les améliorer. Ceci, par la promotion d'activités de réflexion [Perkins et Unger, 1999] durant la lecture sous la forme d'aides et d'incitations à la création de représentations externes exprimant la compréhension de l'apprenant pour les idées présentées dans le document lu.

D'une façon fonctionnelle, un STI basé sur cette approche utilisera le discours exprimé dans le texte lu comme un support d'apprentissage et invitera l'apprenant lors d'une session de lecture, à rendre explicites et d'une manière graphique, *les éléments clés de compréhension*, durant son activité d'annotation ordinaire. Ces éléments de compréhension peuvent être très simples comme la construction ou l'achèvement d'un arbre ontologique ou plus complexes comme l'organisation d'un réseau d'argumentations, de contre-argumentations, de preuves, etc.

Pour que le STI proposé soit capable de suivre et d'analyser le processus de construction de la connaissance réalisé par l'apprenant, il doit contenir une représentation de référence de cette connaissance, que nous appelons le *modèle de référence*. Cette représentation est réalisée par l'expert de la matière (ou le *SME*) et stockée dans la base de connaissance du système. La détection des différences entre les deux représentations, celle de l'apprenant et celle du modèle de référence, peut être le signe d'une mauvaise compréhension ou erreur d'interprétation commise par l'apprenant dans sa lecture du texte. Ainsi, le système est en mesure de donner un feedback immédiat à l'apprenant sur sa représentation et peut, en utilisant le modèle de référence lui donner de l'aide sous forme d'indices ou le référer à l'endroit (ou les endroits) appropriées du texte qu'il doit relire plus attentivement avant de revenir mettre à jour sa modélisation pour la corriger.

Le rôle du concepteur de cours, ou l'expert de la matière, dans un STI basé sur l'approche *LEKC*, est de concevoir à l'aide des outils de conception (*the authoring tools*) les modèles de référence pour les documents didactiques à utiliser dans une formation. On dit qu'on *augmente* les documents didactiques.

Augmenter un document didactique consiste à créer trois modèles de connaissance qui reflètent trois aspects importants dans le contexte de la lecture (Figure 4-1):

- **L'aspect propositionnel**, relatif à la définition des éléments élémentaires du discours, qui sont les propositions, et leur organisation. Cet aspect est modélisé dans *le modèle du discours LEKC-MD*.
- **L'aspect épistémique**, modélisé dans *le modèle épistémique LEKC-ME*, tente de créer des interprétations schématiques (sous forme de graphes) du sens et du raisonnement véhiculés par le discours. Ces formes sont initialement construites par le concepteur de

cours dans le modèle de référence, et reconstruites plus tard par l'apprenant (le lecteur) comme activité de réflexion qui cherche à approfondir sa compréhension du sujet, tout en offrant un moyen de contrôler et de vérifier cette compréhension.

- **L'aspect didactique**, vise à placer l'expérience de lecture dans un cadre de planification didactique global. La lecture didactique est très différente de la lecture récréative. Elle nécessite une planification préalable pour déterminer entre autres, les buts et les objectifs d'une lecture, les documents concernés, les parties sélectionnées ainsi que l'ordre de leur lecture. Le modèle concerné par ces aspects est *le modèle didactique LEKC-MI*.

Dans la suite de ce chapitre nous allons exposer en détail la prise en compte de ces trois modèles dans l'approche *LEKC*. Nous focaliserons davantage sur les fondements théoriques et les aspects formels. Les aspects pratiques relatifs à l'utilisation de notre approche, seront exposés dans les chapitres suivants.

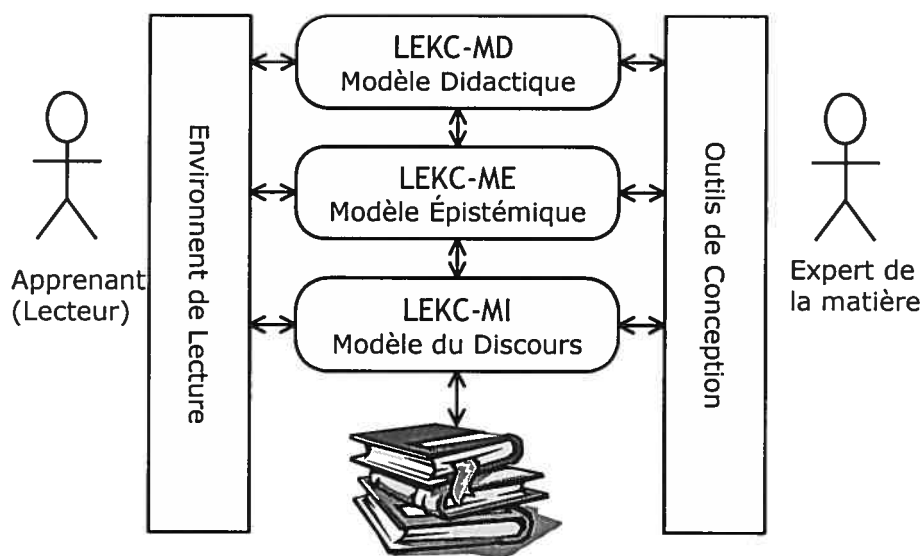


Figure 4-1 – Modèle de connaissance dans l'approche *LEKC*

4.2 Le modèle du discours (*LEKC-MD*)

La recherche en linguistique et en linguistique computationnelle a longtemps indiqué que le texte n'est pas juste une simple séquence de clauses et de phrases, mais plutôt une structure hautement élaborée [Marcu, 1997]. L'unité élémentaire qui constitue le texte a reçu beaucoup

de définitions et de nominations parmi lesquelles on trouve : la phrase [Polanyi, 1988], la clause [Longacre, 1983], l'unité prosodique [Hirschberg et Litman, 1987], le segment de discours [Grosz et Sidner, 1986], la proposition [Denhière et Baudet, 1992] et la microproposition [Dijk et Kintsch, 1983]. Mais malgré ces différences dans la définition et la nomination de ces éléments élémentaires du texte, un consensus se dégage sur le fait que ces éléments ne se chevauchent pas et qu'il existe des relations de plusieurs natures entre ces éléments, comme les relations grammaticales par exemple, et les relations rhétoriques (e.g. *Elaboration, But, ...etc.*).

Le modèle du discours dans l'approche *LEKC*, noté *LEKC-MD*, a pour premier objectif de supporter l'activité d'annotation entreprise par l'apprenant dans la lecture des documents didactiques, spécialement, les opérations de sélection (le marquage), de condensation et d'abstraction d'idées. Comme ces activités, telles qu'elles sont pratiquées, ne concernent généralement que les éléments importants du texte lu, le modèle du discours ne cherche pas la complétude et ne traite qu'un sous-ensemble des éléments du discours (ceux jugés importants par le concepteur). Le modèle du discours est donc caractérisé par une grande parcimonie.

4.2.1 La dimension propositionnelle

Le modèle du discours concerne la représentation du discours véhiculé dans le texte sur le plan des propositions et de leur hiérarchie. Ce modèle est basé sur la théorie de compréhension de discours de Dijk et Kintsch [Dijk et Kintsch, 1983]. Cette théorie stipule que comprendre un texte lors de sa lecture, revient à construire mentalement une séquence ordonnée de propositions et à établir des liens entre ces propositions de manière à aboutir à une hiérarchie de propositions. Ce processus est une approche répétitive faisant intervenir des traitements cognitifs d'abstraction et de généralisation d'idées. Cette construction s'élabore à trois niveaux [Dijk et Kintsch, 1983] :

1. Le niveau des micropropositions : correspond au niveau élémentaire du discours et représente le discours comme un ensemble de micropropositions communiquant un élément de compréhension élémentaire, dit unité cognitive.
2. Le niveau des macropropositions : correspond à un niveau de compilation et d'abstraction des éléments présents dans le niveau précédent (des

micropropositions), par application de règles sémantiques comme la suppression, la généralisation et la construction.

3. Le niveau des superstructures schématiques : correspond à un niveau organisationnel global qui fait que lors de la lecture d'une thèse par exemple on s'attend à une structure globale particulière qui est différente de la structure d'un roman par exemple.

En se basant sur une adaptation des deux premiers niveaux de cette théorie, nous définissons le modèle du discours *LEKC-MD* d'un document, comme étant un sous-ensemble de l'ensemble des micropropositions et des macropropositions contenues dans son discours.

4.2.2 Les micropropositions

La structuration en micropropositions dans le modèle du discours concerne l'organisation de surface du discours. À ce niveau le texte n'est pas interprété. C'est une opération de sélection et aucune modification ou transformation du texte n'est introduite.

Les micropropositions sont de deux types, atomiques et composées:

1. **Les micropropositions atomiques ($m\mathcal{A}$)** : sont des structures de base qui correspondent à un segment de texte. Leur nature linguistique est libre et peut correspondre à un mot, une suite de mots, une phrase ou une suite de phrases. La seule contrainte à respecter est la contiguïté (c.à.d. le texte concerné doit se suivre sur le document). Le système les traite comme des éléments unitaires et garde en permanence pour chacune d'elles, une référence vers sa position exacte sur le document source. La Figure 4-2 donne un exemple de décomposition en micropropositions d'un texte. Les micropropositions (A-“Elle a assuré sa qualification”) et (B-“Bulgarie”) sont considérées comme atomiques.
2. **Les micropropositions composées ($m\mathcal{C}$)** : sont des structures formées par la combinaison de micropropositions atomiques et/ou des micropropositions composées. Les éléments de cette composition n'ont pas besoin de référer à des segments de texte qui sont contigus, comme c'est le cas des micropropositions atomiques.

Les micropropositions composées permettent de référer à un groupe de segments de

texte en tant qu'une seule unité. Le texte qui représente une microproposition composée est formé par la concaténation des textes des micropropositions qui la composent. Si ces micropropositions ne réfèrent pas à des segments contigus, des points de suspension “...” sont insérés pour indiquer cette non-contiguïté. C'est le cas de la microproposition (D) sur la Figure 4-2, qui a pour texte “Elle a assuré sa qualification ... par deux buts à zéro”.

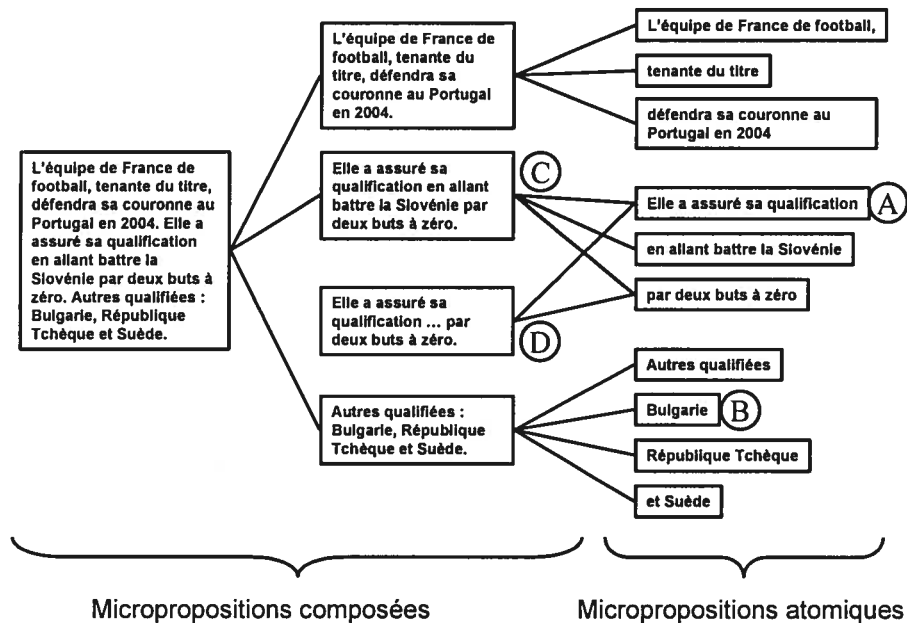


Figure 4-2 – Exemple de décomposition en micropropositions

La création des micropropositions permet d'identifier et sélectionner les segments du discours (e.g. mots, phrases et groupe de phrase) qui vont servir comme briques de travail pour les autres activités de modélisation que nous allons décrire par la suite. Sur le plan pratique, les éléments ainsi identifiés deviennent des éléments actifs sur le document source et peuvent recevoir les actions de l'utilisateur, comme le pointage et la sélection avec la souris. Nous expliquerons les aspects techniques de cette opération lors de la présentation de l'implantation du système.

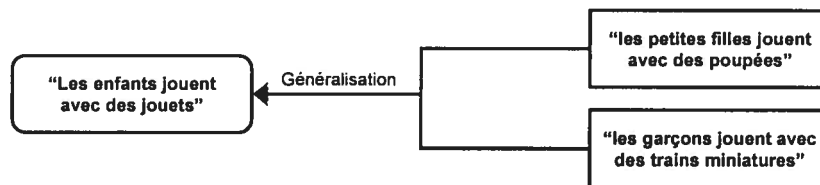
4.2.3 Les macropropositions

Pour faire abstraction des détails lors de la lecture et focaliser sur ce qui est essentiel, le lecteur a recours à des traitements cognitifs visant la généralisation et la condensation de

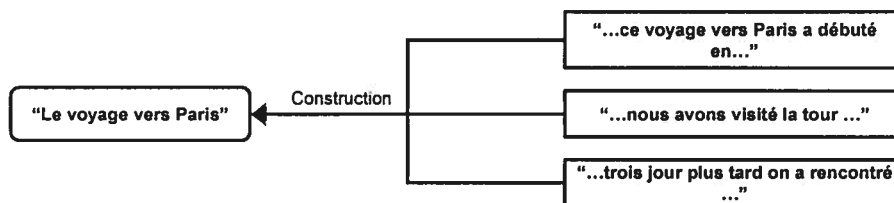
l'information sémantique contenue dans le texte [Denhière, 1985]. À travers ces traitements cognitifs, les micropropositions du discours, sont successivement réduites dans un processus itératif pour produire des macropropositions exprimant un sens de plus en plus global [Dijk et Kintsch, 1983]. À l'inverse des micropropositions, les macropropositions sont une *interprétation* du texte.

Les traitements cognitifs sont de trois natures : la *suppression*, la *généralisation* et la *construction*.

- **La suppression** (*Deletion*) : est le traitement cognitif par lequel les propositions qui ne sont pas une condition directe ou indirecte pour l'interprétation d'autres propositions, sont ignorées. Par exemple les descriptions détaillées, les informations d'arrière plan, les connaissances communes, sont généralement ignorées une fois lues.
- **La généralisation** (*Generalization*) : est le traitement cognitif par lequel une séquence de propositions peut être remplacée par une seule proposition qui est une généralisation de cette séquence de propositions. C'est la construction d'une proposition générale à partir du détail sémantique d'un groupe de propositions. Par exemple, à partir des propositions "les petites filles jouent avec des poupées" et "les garçons jouent avec des trains miniatures" peuvent être généralisées en une seule proposition "Les enfants jouent avec des jouets".



- **La construction** (*Construction*) : est le traitement cognitif par lequel une séquence de propositions peut être remplacée par une seule proposition qui est une implication de l'ensemble joint des éléments de cette séquence de propositions. Par exemple, une séquence de propositions décrivant tous les petits détails d'un voyage peut être ramenée à une seule proposition indiquant que ce voyage à eu lieu en faisant abstraction des détails, si ces détails ne sont pas nécessaires pour la compréhension du reste de l'histoire.



Les macropropositions dans le modèle du discours sont des propositions liées par des relations de type *Généralisation* ou *Construction* avec d'autres macropropositions ou micropropositions. Le texte d'une macroproposition est une interprétation résumée de l'information sémantique présente dans l'ensemble des propositions qui lui sont rattachées. C'est une forme de *raisonnement par déduction* basé sur ces propositions. Cette activité est une pratique courante dans le processus d'annotation conventionnelle lors de la lecture. On retrouve sa manifestation dans la création de paraphrases, généralement sur les marges des livres ou sur des cahiers de notes.

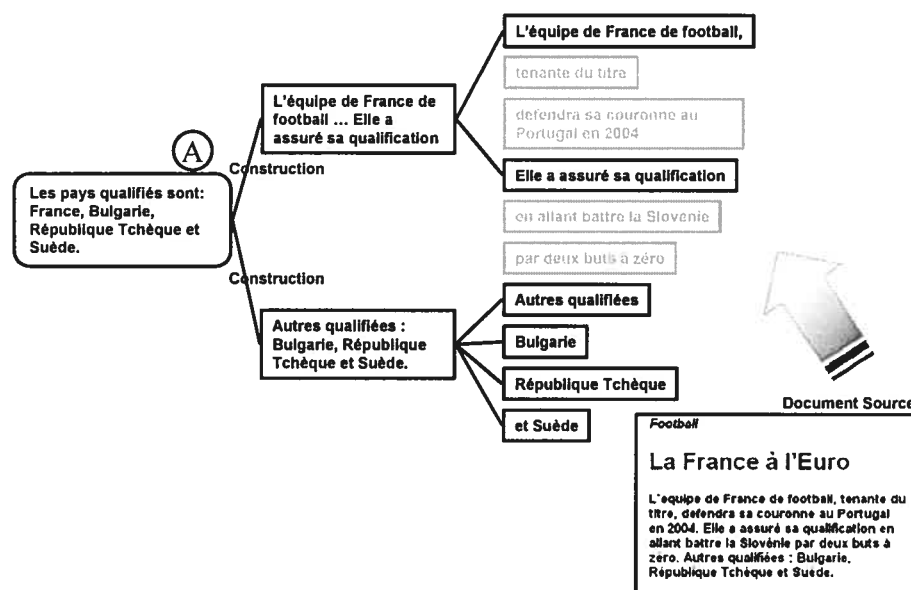


Figure 4-3 – Exemple de construction de macroproposition

La création par le concepteur du cours des macrostructures, ainsi que leurs relations de *Généralisation* et de *Construction*, produit une structure de connaissance qui permet au lecteur d'avoir des éléments d'annotation sous forme de propositions *condensées* prêtes à l'utilisation. De plus le système est capable, en exploitant le réseau des relations de *Généralisation* et *Construction*, d'expliquer au lecteur (l'apprenant) l'origine de ces propositions. L'apprenant

peut alors suivre et d'une façon graduelle la succession des *déductions* qui ont abouti à une macroproposition donnée, jusqu'à son origine dans le document source. La Figure 4-3-(A) donne un exemple simplifié d'une macroproposition avec cette chaîne de déductions.

4.2.4 Modélisation formelle du modèle du discours

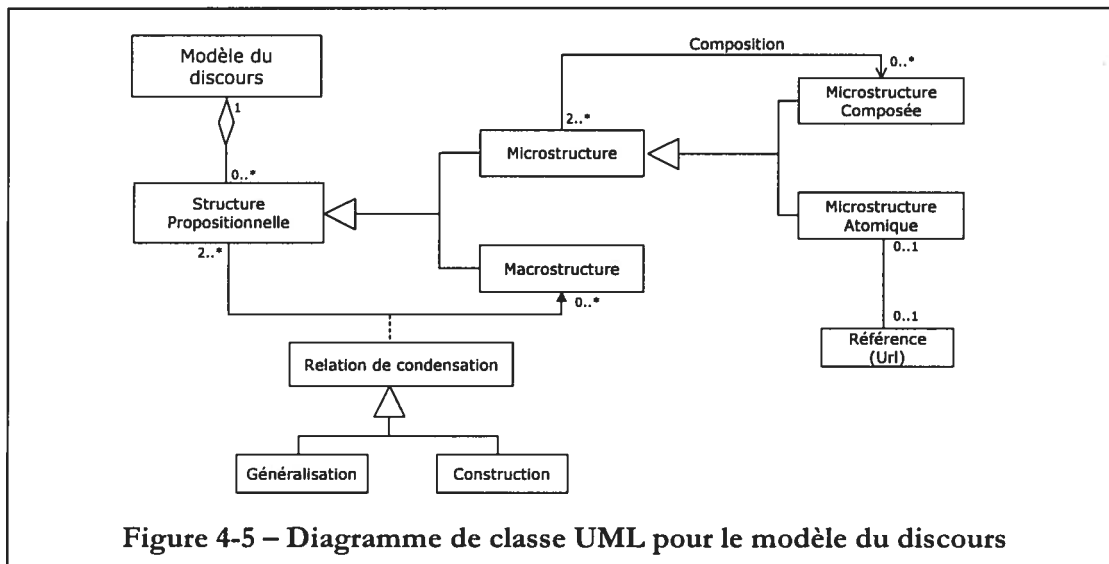
Le modèle du discours peut être défini et formellement en utilisant la notation BNF (Backus Naur Form¹) comme indiqué dans Figure 4-4.

| | |
|-----------------------------|---|
| <ModèleDuDiscours> | ::= <StructurePropositionnelle>* |
| <StructurePropositionnelle> | ::= <Microproposition> <Macroproposition> |
| <Microproposition> | ::= <MicropropositionAtomique> <MicropropositionComposée> |
| <MicropropositionAtomique> | ::= <SegmentDeTexte> <UrlRéférence> |
| <SegmentDeTexte> | ::= <Texte> |
| <UrlRéférence> | ::= <Texte> |
| <MicropropositionComposée> | ::= <Microproposition> <Microproposition>+ |
| <Macroproposition> | ::= <Proposition><RelationCondensation><StructurePropositionnelle>+ |
| <Proposition> | ::= <Texte> |
| <RelationCondensation> | ::= Construction Généralisation |

Figure 4-4 - Définition du modèle du discours en notation BNF

Dans cette définition le modèle du discours est un graphe formé d'un ensemble de structures propositionnelles (microstructures et macrostructures) reliées par des relations de composition, de généralisation et de construction. Une structure propositionnelle est soit une microproposition ou une macroproposition. Une microproposition peut être atomique ou composée. Une microproposition atomique est un segment de texte avec une référence pointant sur sa localisation dans le document source. Une microproposition composée est une proposition qui a au minimum deux relation de composition avec d'autres micropropositions. Une macroproposition est une proposition avec au moins une relation de condensation (généralisation ou construction) avec d'autres structures propositionnelles. La Figure 4-5 donne une modélisation UML de cette définition.

¹ Voir annexe A



4.2.5 Utilité du modèle du discours

L'application des règles sémantiques pour la production des macropropositions est un processus de transformation qui simplifie le document lu et lui procure une organisation structurelle supplémentaire qui se manifeste sous la forme d'une hiérarchie propositionnelle. À la base de cette hiérarchie on retrouve les segments du texte source (Figure 4-6). Les idées brutes et détaillées de ces segments de texte sont successivement combinées et réduites pour produire des idées (des propositions) plus générales et plus abstraites, à mesure qu'on monte dans cette hiérarchie de propositions.

Le modèle du discours est destiné à jouer plusieurs rôles dans la lecture d'un document didactique. Premièrement, il donne au lecteur (l'apprenant) une vision et une perspective supplémentaires sur les idées exprimées dans le document lu ainsi que leur organisation. Deuxièmement, il devient un support efficace pour l'annotation. En plus de pouvoir marquer (*highlight*) des segments de texte comme dans l'annotation conventionnelle, le lecteur a aussi la possibilité de marquer les propositions qui l'intéressent directement sur la hiérarchie propositionnelle au niveau de détail qui le satisfait (plus détaillé à la base de la hiérarchie et plus concis vers les sommets). Et troisièmement, il fournit les briques de construction pour les modèles de connaissances plus évolués que nous allons explorer par la suite.

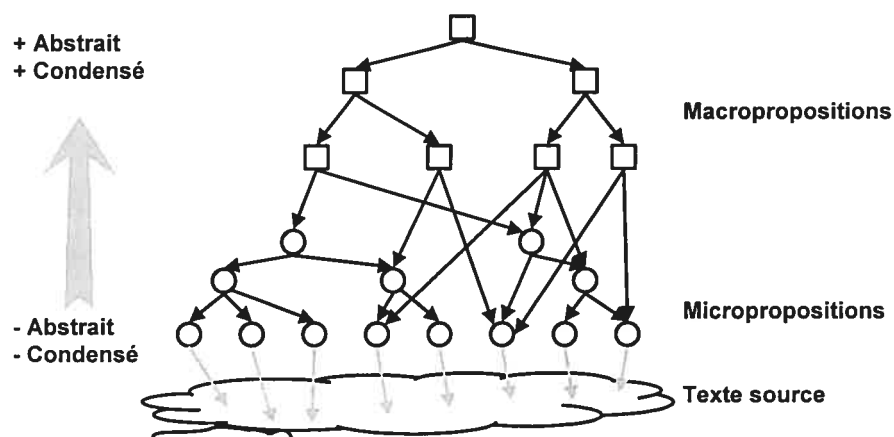


Figure 4-6 – L’abstraction et généralisation dans le modèle du discours

Ce modèle peut être suffisant dans bien des cas, seulement, considéré dans le contexte d’usage que nous voulons en faire, il souffre d’une lacune importante. Il ne prend pas en considération la multitude de relations qui peuvent exister entre les propositions et dont la mise en évidence peut grandement influencer la compréhension. Par exemple, des relations de *cause-à-effet* ou de *condition* entre propositions, bien qu’elles soient très importantes à la compréhension du contexte, ne sont pas prises en compte dans le modèle du discours. C’est pour cette raison que nous introduisons le modèle épistémique *LEKC-ME*.

4.3 Le modèle épistémique (*LEKC-ME*)

La compréhension du discours n’est pas juste la construction d’un réseau de propositions produites par le traitement cognitif de chaque phrase, mais elle est la construction d’un modèle mental qui est progressivement révisé et enrichi, et où les éléments du discours agissent comme des instructions à exécuter pour mettre à jour et faire évoluer ce modèle mental [Denhière et Baudet, 1992].

Nous convenons bien que ce processus cognitif soit largement interne (au lecteur), néanmoins, dans une lecture réfléchie comme c’est le cas dans la lecture didactique, le lecteur a tendance à extérioriser une partie de ce processus à travers son activité d’annotation avancée, outre la sélection et le marquage, sous la forme de *représentations externes* [Cox et Brna, 1995].

Les représentations externes peuvent prendre des formes très variées. Elles peuvent être simples, comme une liste d'éléments que le lecteur a dégagés du texte lu, ou elles peuvent être complexes, comme par exemple des tableaux ou des figures. Dans notre modèle nous nous intéressons aux formes en graphe, plus générales et plus flexibles. Nous les avons nommées *formes épistémiques* [Rouane, Frasson et Kaltenbach, 2003-a].

Une *forme épistémique* est une représentation graphique d'une structure de connaissance cible, construite par le concepteur de cours et reconstruite plus tard par l'apprenant (le lecteur) lors de la lecture, et qui montre l'organisation de la connaissance et illustre les relations qui existent entre les différents faits et concepts énoncés dans le discours d'un document [Rouane, Frasson et Kaltenbach, 2003-a]. Une *forme épistémique* est un graphe $\langle N, R \rangle$ où :

- N est un ensemble de nœuds propositions qui peuvent être des micropropositions, des macropropositions, ou des propositions thématiques (telles *le concept, l'affirmation, la preuve*, etc.) Ces notions seront définies en détail un peu plus loin.
- R est un ensemble de relations de type rhétorique (tel *But, Moyen, Condition,...*) ou de type thématique (tel *Affaiblit, Renforce, Contredit,...*), entre deux nœuds de N, deux relations de R ou entre un nœud de N et une relation de R.

4.3.1 Les dimensions de la modélisation épistémique

Dans le cadre d'un support à l'annotation lors de la lecture, nous considérons les formes épistémiques sous trois dimensions : *la dimension rhétorique, la dimension thématique, et la dimension structurelle*.

La dimension rhétorique tente de mettre en évidence les relations rhétoriques entre les éléments d'un discours. Ces relations sont du genre *But, Moyen, Condition*, etc. Elles s'appliquent généralement à des nœuds de type microproposition ou macroproposition et montrent les rôles que jouent les éléments du discours les uns vis-à-vis des autres. Ce qui est très important pour la compréhension du sens d'un discours.

La dimension thématique tente de mettre en évidence les idées ou thèmes essentiels d'un discours scientifique. C'est une analyse qui va plus loin que la rhétorique et cherche à identifier et mettre en évidence des notions comme les *concepts* référés dans le discours, les

problèmes traités, les *solutions* apportées, avec des relations comme *affaiblir*, *renforcer*, *supporter*, *contredire*, etc. Des notions qu'on trouve très souvent dans les annotations des étudiants.

La **dimension structurelle** tente de capter les aspects structurels présents dans le discours. Elle traite des aspects comme la notion de *liste* ou de *séquence* qui sont très utilisés dans l'activité d'annotation pour énumérer des objets ou des événements, et des aspects ontologiques qui sont souvent représentés par des relations entre concepts du genre *est-une(e)* ou *est-membre-de*.

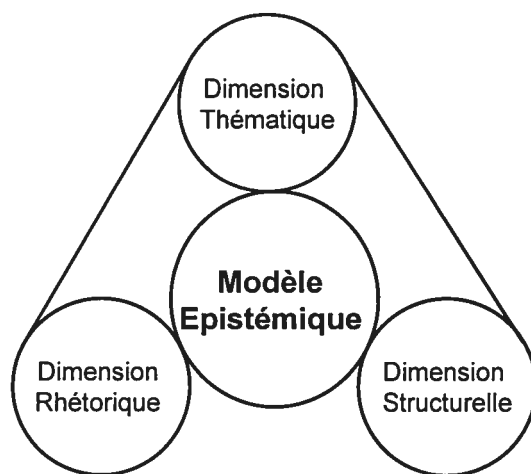


Figure 4-7 – Couplage des différentes dimensions

Le modèle épistémique *LEKC-ME* est défini comme le couplage de ces trois dimensions (Figure 4-7). La définition détaillée pour chacune de ces dimensions est donnée dans la suite de ce chapitre. Ces définitions ne sont pas restrictives. Des extensions à ces définitions peuvent être faites par le concepteur de cours quand il le juge nécessaire. D'un autre côté il ne faut pas considérer ces dimensions comme des couches de modélisation qui doivent se succéder. La modélisation épistémique est un mélange de toutes ces dimensions.

Il est important de noter que les définitions données ne cherchent nullement l'exhaustivité ou la rigueur absolue. Les différents types de nœuds et de liens proposés ne sont donnés qu'à titre indicatif pour montrer aux utilisateurs du système (concepteurs de cours et apprenants) le genre de connaissance à mettre au niveau de chaque dimension. La recherche de la simplicité est volontaire dans ce contexte pour faciliter la manipulation de la connaissance et éviter qu'elle ne devienne une distraction ou même un frein à l'apprentissage.

4.3.2 La dimension rhétorique

La dimension rhétorique dans l'approche *LEKC*, tente de tenir compte des relations rhétoriques qui existent entre les éléments propositionnels du discours (définis dans le modèle du discours). Nous l'avons basée sur la théorie de la structure rhétorique ou la RST (Rhetorical Structure Theory). La RST a initialement été conçue dans le cadre des études sur la génération automatique du texte [Mann et Thompson, 1988], et est devenue par la suite la théorie d'analyse du discours la plus dominante durant la dernière décennie.

Au centre de cette théorie on trouve la notion de la relation rhétorique : une relation qui peut exister entre deux segments d'un texte nommés le noyau N (*nucleus*) et le satellite S (*satellite*). Le noyau N, comparativement au satellite S, est le segment du texte supposé exprimer ce qui est le plus important pour le but de l'auteur. La relation rhétorique exprime le rôle joué par N et S l'une vis-à-vis de l'autre.

Le processus d'analyse dans la RST, a pour objectif de déterminer une partie de ce qui soutient la compréhension par l'identification des relations rhétoriques qui existent entre les segments importants d'un texte. La modélisation selon la dimension rhétorique retient le même principe de relations rhétoriques mais avec son extension aux formes épistémiques. En d'autres termes, le noyau N et le satellite S ne sont plus limités à des segments de textes (des micropropositions) mais peuvent être eux-mêmes des formes épistémiques. Et dans cette optique, les relations rhétoriques dans la dimension rhétorique expriment les rôles que jouent les formes épistémiques les unes par rapport aux autres (Figure 4-8).



Figure 4-8 – Illustration d'une relation rhétorique de justification entre deux formes épistémiques

4.3.2.1 Exemple de modélisation rhétorique.

Pour montrer l'intérêt de la modélisation rhétorique dans l'approche *LEKC* considérons l'analyse du texte simple suivant qui parle de *depsydres* :

Les clepsydes La clepsydre est une horloge à eau connue aussi bien des Egyptiens que des Amérindiens ou que des Grecs. Un vase percé d'un trou laisse couler de l'eau. Des graduations situées à l'intérieur permettent de mesurer des intervalles de temps. La clepsydre a une forme évasée, plus large en haut, car le débit de l'eau est plus grand quand la dénivellation est plus grande. Les graduations sont ici à peu près équidistantes. Si le cadran solaire donne l'heure pendant le jour, la clepsydre fait la même chose la nuit, et elle mesure en plus des durées plus brèves avec une bonne précision. ...

A ce stade de la modélisation, nous supposons que les éléments du discours sont déjà segmentés dans le modèle du discours *LEKC-MD*, en micropropositions et macropropositions. Pour la modélisation rhétorique, le concepteur de cours crée des formes épistémiques (structures en graphe) en plaçant des propositions et en indiquant les relations rhétoriques appropriées qui expriment leurs rôles les unes vis-à-vis des autres. Le résultat de cette modélisation peut ressembler aux formes épistémiques données dans la Figure 4-9 et la Figure 4-10. Le concepteur peut créer deux genres de formes :

- Des formes étalées où un nœud qui représente une forme épistémique est mis dans le même graphe. Ceci est pour les cas simples qui n'encombrent pas la représentation, comme dans le cas (A) sur la Figure 4-9.
- Des formes condensées où un nœud qui représente une forme épistémique est remplacé par une proposition à laquelle on associe un hyperlien vers cette forme épistémique qui est la modélisation rhétorique de cette proposition. C'est le cas dans l'exemple (B) sur la Figure 4-9 qui renvoie sur la représentation rhétorique de la Figure 4-10.

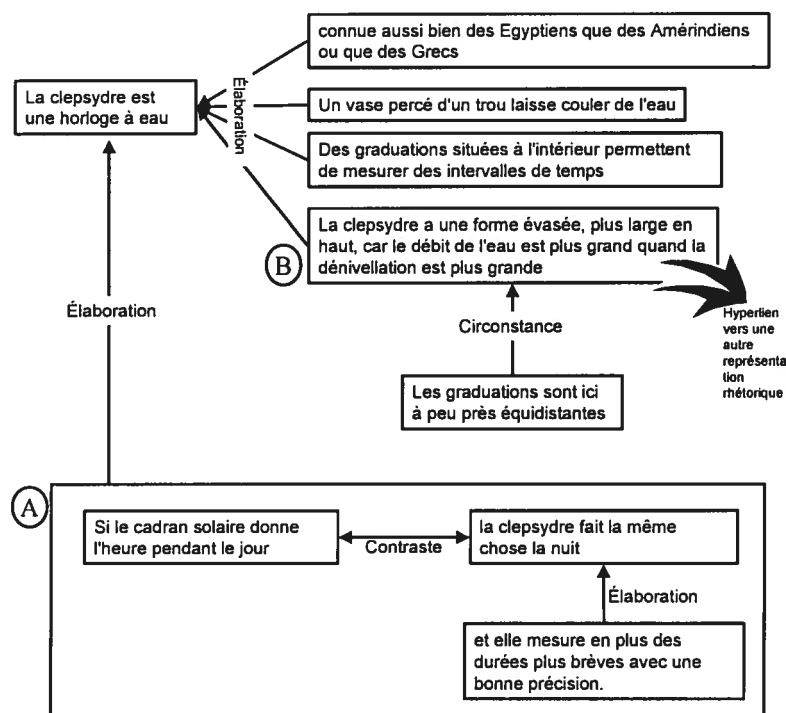


Figure 4-9 – Exemple simple d'analyse rhétorique (1)

Le résultat de cette modélisation rhétorique peut être exploité de plusieurs façons dans l'approche *LEKC* :

- Il fournit au concepteur une nouvelle perspective synthétique du discours contenu dans le texte. Avec cette aide, il peut relever les lacunes dans le discours et peut y remédier par ajout d'explications ou, quand cela est possible, par modification du texte source initial.
- Le résultat de la modélisation peut être présenté à l'apprenant comme matériel didactique supplémentaire au texte et qui peut servir même imprimé sur un support papier. L'activité intellectuelle de faire la correspondance (le *mapping*) entre le texte source et ces représentations, grâce à la réflexion qu'elle nécessite, peut aider à approfondir la compréhension.
- Le *module tuteur* du système *LEKC*, que nous allons présenter dans les prochains chapitres, et qui est responsable de superviser l'activité de lecture, utilise ces représentations pour générer des activités de réflexion lors de la lecture, comme

demander au lecteur d'identifier certaines des relations rhétoriques par exemple, ou de remettre en ordre une représentation désordonnée, ou de retrouver dans le texte des propositions manquantes à une représentation.

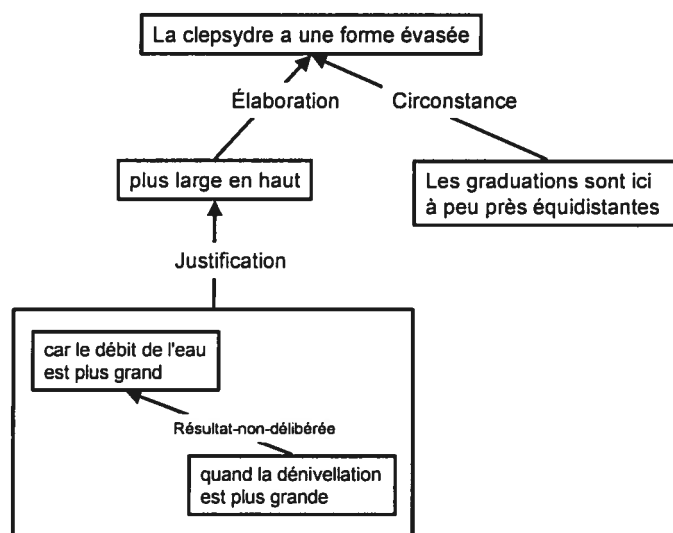


Figure 4-10 – Exemple simple d'analyse rhétorique (2)

4.3.2.2 Les relations de la dimension rhétorique

La modélisation rhétorique d'un document didactique est basée sur l'identification des relations rhétoriques importantes qui relient les éléments du discours et dont la mise en évidence peut aider à la compréhension du sens véhiculé. Les relations rhétoriques sont de deux types: les *relations de présentation* et les *relations du domaine*.

Les relations rhétoriques de présentation

Une relation de présentation entre deux formes épistémiques N et S, est une relation qui a pour effet d'accroître l'inclination du lecteur pour supporter une position donnée sur N à la lecture de S, comme par exemple son degré de jugement positif envers N ou son acceptation de celle-ci. On dit que ces relations sont *informationnelles*. Considérons l'exemple suivant :



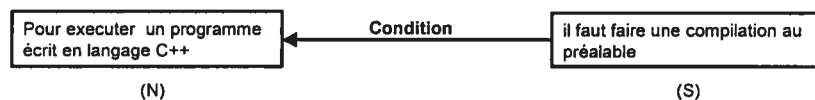
Dans cet exemple, S est lié à N par une relation de justification et a pour effet d'accroître notre acceptation de N. Notons qu'une proposition est une forme épistémique formée de cette seule proposition, et que par conséquent, N et S sont dans cet exemple sont des formes épistémiques.

La liste suivante donne la définition de l'ensemble des relations rhétoriques de présentation retenues pour la modélisation dans *LEKC-ME*. Dans ces définitions on suppose que ces relations sont entre deux formes épistémiques, l'une jouant le rôle du noyau N et l'autre jouant le rôle du satellite S.

- **Antithèse** : les formes épistémiques N et S sont en contraste et on ne peut pas avoir un regard positif sur les deux en même temps. S est jugée négativement, ce qui augmente notre position positive envers N.
- **Concession** : N est jugée positivement malgré l'incompatibilité potentielle ou apparente avec S.
- **Arrière-plan** : S sert à faciliter la compréhension de N et donne un arrière-plan pour cette compréhension.
- **Préparation** : S présente un texte préparant le lecteur à anticiper et à interpréter le texte qui va être présenté dans N.
- **Facilitation** : N présente une action et S fournit une information destinée à aider le lecteur à accomplir cette action.
- **Démonstration** : N présente une affirmation et S présente une information pour accroître la croyance du lecteur dans N.
- **Justification** : S présente une information légitimant et justifiant N.
- **Motivation** : S présente une information destinée à accroître le désir du lecteur pour accomplir l'action décrite par N.
- **Reformulation** : S est une reformulation de N. S et N peuvent être de taille comparable mais N est considérée plus centrale pour le but de l'auteur.
- **Résumé** : comme dans la relation de reformulation, S est une reformulation de N, seulement S à une taille réduite comparativement à N.

Les relations rhétoriques du domaine

Une relation du domaine entre les formes épistémiques N et S est une relation dont l'effet souhaité est que le lecteur reconnaisse consciemment l'existence de cette relation. Elle a un sens beaucoup plus fort qu'une relation de présentation et elle est plus explicite. On dit que ces relations sont *intentionnelles*. Considérons l'exemple suivant :



Dans cet exemple S est lié à N par une relation de condition et l'auteur cherche à rendre le lecteur parfaitement conscient de l'existence de cette condition.

La liste suivante donne la définition de l'ensemble des relations rhétoriques du domaine retenues pour la modélisation dans *LEKC-ME*. Dans ces définitions on suppose que ces relations sont entre deux formes épistémiques, le noyau N et le satellite S.

- **But** : N décrit une activité et S décrit une situation non réalisée. Le lecteur reconnaît que N est entreprise pour réaliser S.
- **Cause-délibérée** : N décrit une action ou une situation qui découle d'une action délibérée, et S décrit une situation. S est causée par N du fait de l'action délibérée de quelqu'un.
- **Cause-non-délibérée** : N et S décrivent deux situations. S est causée par N, mais pas du fait de l'action délibérée de quelqu'un.
- **Circonstance** : S définit un cadre interprétatif temporel ou situationnel dans lequel le lecteur doit interpréter N.
- **Évaluation** : S donne une évaluation de N.
- **Interprétation** : S donne une interprétation de N.
- **Élaboration** : S procure de l'information supplémentaire pour N qui contient des informations de base.
- **Moyen** : S décrit la méthode ou l'outil utilisé qui rend possible la réalisation de N.

- **Résultat-délibéré** : S est une situation résultante de N du fait de l'action délibérée de quelqu'un.
- **Résultat-non-délibéré** : S est une situation résultante de N, mais pas du fait de l'action délibérée de quelqu'un.
- **Solution** : S présente un problème auquel N apporte une solution totale ou partielle.
- **Condition** : S est une situation conditionnante pour N. Le lecteur doit comprendre que la réalisation de N dépend de la réalisation de S.
- **Anti-condition** : S est une situation conditionnante pour N. N est réalisable si S n'est pas réalisée.
- **Inconditionnel** : cette relation exprime l'inverse de ce qu'exprime la relation Condition. Elle indique que S n'est pas une situation conditionnante pour N. Le lecteur doit comprendre que la réalisation de N ne dépend pas de la réalisation de S.
- **À-moins-que** : S décrit une situation qui affecte N. N est réalisable pourvu que S ne le soit pas.

4.3.2.3 Différence avec la RST classique

Le processus d'analyse dans la RST a pour objectif de déterminer la structure rhétorique pour un document donné dans sa totalité. Un exemple de ce genre d'analyse est donné dans la Figure 4-11.

[Mann et Thompson, 1988] dans leur définition de la RST ont posé quatre contraintes sur la construction de la structure rhétorique d'un discours :

- La complétude (*Completeness*) : le niveau le plus supérieur de la construction (la racine) couvre le texte en entier.
- La connectivité (*Connectedness*) : excepté la racine, chaque segment du texte est soit un noyau, un satellite ou un constituant d'une structure composée
- L'unicité (*Uniqueness*) : chaque structure est constituée de segments de texte différents et chaque relation à l'intérieur d'une structure s'applique à des segments de texte différents.

- L'adjacence (*Adjacency*) : les segments de texte d'une structure forment un segment contigu s'ils sont réunis.

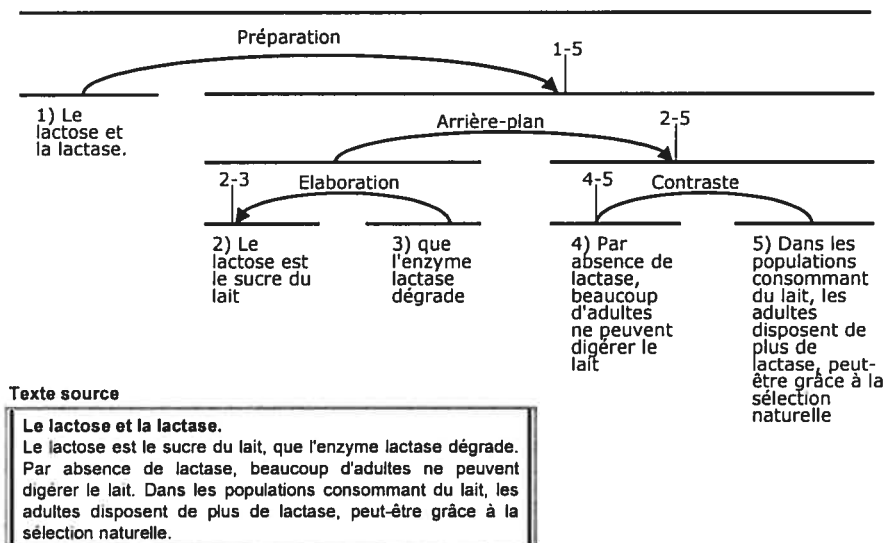


Figure 4-11 – Exemple d'une analyse rhétorique complète [Mann, 2003]

À ces contraintes explicites exprimées par [Mann et Thompson, 1988], nous ajoutons deux autres contraintes implicites qui découlent directement de celles-ci :

- La base textuelle : les éléments de base dans cette analyse sont les segments de texte.
- La structuration en arbre (*Tree structure*) : la structure construite est toujours sous la forme d'un arbre.

Ces contraintes prennent tout leur sens dans le contexte d'analyse ou de génération automatique de texte, mais ils sont inappropriés dans le contexte d'un support à l'annotation vue les difficultés et l'effort qu'ils introduisent dans la conception. Ce qui nécessite l'introduction de relaxations sur ces contraintes.

La complétude (*Completeness*) n'est pas recherchée dans notre modélisation du fait que seules les notions importantes et qui peuvent vraiment aider à la compréhension du discours sont représentées. De ce fait, notre modèle de la rhétorique est nécessairement fragmentaire et incomplet.

La connectivité (*Connectedness*) telle quelle est définie n'est pas obligatoire dans notre modèle. La modélisation peut se faire plus facilement avec un ensemble de petites structures de

représentation (en graphe) non connectées, qu'avec une seule grande structure qui peut rapidement devenir difficile à gérer et à visualiser.

L'unicité (*Uniqueness*) reste valable dans notre modèle mais par contre l'adjacence (*Adjacency*) ne l'est pas. Les segments de texte utilisés dans la modélisation d'une structure n'ont pas besoin d'être contigus dans le texte source. Ce qui donne plus de flexibilité dans le choix de ces segments.

L'introduction de ces relaxations sur les contraintes originales de la RST fait en sorte que la structure de représentation rhétorique dans l'approche *LEKC* n'est plus un arbre mais un graphe qui peut être non connexe. D'autre part les éléments de base ne se limitent plus aux segments de texte mais peuvent être des formes épistémiques allant des plus simples comme les micropropositions, au plus compliquées comme les formes en graphes. Des structures qui sont plus élaborées que les segments du texte et qui ne sont pas directement disponibles dans le texte.

4.3.2.4 Limites de la modélisation rhétorique.

La dimension rhétorique retient l'essentiel d'un discours sur le plan rhétorique. Il exprime les relations entre les éléments du discours et détermine par la même occasion, les rôles qui jouent ces éléments les uns vis-à-vis des autres. Clarifier et rendre explicite ce genre de rôles et de relations facilite et canalise la compréhension du lecteur. Seulement, cela n'est pas suffisant pour capter toute la complexité d'un contexte de lecture. Cette dimension ne tient pas en compte des notions importantes relatives aux thèmes généraux traités dans les discours didactiques, comme la notion du *concept* ou la notion du *problème*. Nous tentons de remédier à cette situation par l'introduction de la dimension thématique.

4.3.3 La dimension thématique

La dimension thématique dans l'approche *LEKC* tente de représenter les idées essentielles ou thèmes d'un texte didactique scientifique. C'est une analyse plus abstraite comparativement à l'analyse rhétorique qui est très liée au discours et ses éléments. À ce niveau, nous cherchons à identifier et à mettre en évidence des notions comme les *questions* soulevées dans un texte, les *hypothèses*, les *preuves*, les *concepts* utilisés, etc.

4.3.3.1 Les types thématiques

La dimension thématique fait introduire dans la modélisation d'une forme épistémique un certain nombre de nœuds et de liens thématiques. Un nœud thématique (ou simplement thème) est une proposition à laquelle on associe un *type thématique*. Les type thématiques prédéfinis dans le système sont :

- **L’Affirmation** : c’est une proposition exprimant une position à propos d’un problème ou d’une situation. Elle peut être de quatre natures :
 - 1) **La Conjecture**, une affirmation avec peu ou pas de preuves.
 - 2) **L’Hypothèse**, une affirmation qui est dérivée d’un théorème ou basée sur des preuves empiriques et expérimentales.
 - 3) **Le Fait**, une affirmation bien supportée par des preuves empiriques et expérimentales.
 - 4) **L’Axiome**, une affirmation généralement acceptée comme vraie et qui, de ce fait, peut être utilisée comme base d’inférence ou d’argumentation.
- **Le Concept** : c’est une construction élémentaire pour symboliser des notions concrètes ou abstraites du domaine. Il n’y a pas de définition stricte et exacte pour la notion de concept, mais une caractéristique essentielle qu’il doit avoir est l’atomicité.
- **La Preuve** : ce sont les données rassemblées dans le but de supporter ou de réfuter une affirmation donnée. Les preuves peuvent être qualitatives ou quantitatives.
- **La Méthode** : une proposition décrivant ou référant à une procédure ou une technique utilisée pour générer des preuves pour une affirmation donnée ou pour résoudre un problème.
- **Le Problème** : une proposition décrivant ou référant à une situation ou à un sujet qui pose des difficultés et qui demande à être solutionné ou contrôlé.
- **La Question** : un sujet, un point ou un problème qui donne lieu à réflexion et dont une réponse doit être suggérée (question fermée) ou non (question ouverte).
- **La Théorie** : une formulation systématique à propos d’un problème particulier du domaine, dérivable par des procédures de déduction et d’induction.

4.3.3.2 Les relations thématiques

Les thèmes sont reliés par des relations thématiques ou des relations rhétoriques (définies précédemment). La liste des relations thématiques prédéfinies dans le système est donnée dans ce qui suit :

- **Affaiblit / Renforce** : une paire de relation pour indiquer la notion d'influence positive (*renforce*) ou négative (*affaiblit*) entre les deux éléments en relation. Exemple, une affirmation qui peut affaiblir (ou renforcer) une certaine théorie.
- **Supporte / Contredit** : une paire de relation pour indiquer la notion de support ou de contradiction entre les deux éléments en relation. Exemple, une preuve qui supporte (ou contredit) une certaine affirmation.
- **Alternative-à** : est une relation qui indique la notion d'équivalence (ou d'interchangeabilité) entre les deux éléments en relation.
- **Répond-à** : cette relation entre un élément A et un élément B indique que A apporte un élément de réponse à B qui est généralement un problème ou une question.
- **Réfère-à** : cette relation entre un élément A et un élément B indique que l'information contenue dans A réfère à l'élément B. B est le plus souvent un concept.
- **Suggère** : cette relation entre un élément A et un élément B indique que B est suggéré par A.

4.3.3.3 Exemple d'analyse thématique

L'exemple suivant illustre l'utilisation de l'analyse thématique dans l'approche *LEKC* (Figure 4-12). Nous avons choisi un segment d'un texte scientifique qui traite de la présence de l'eau sur la planète Mars

À ce stade nous supposons que le modèle du discours (découpage en micropropositions et macropropositions) est déjà réalisé. La modélisation rhétorique peut ou pas être faite. Le concepteur du cours peut alors décider de ce qui est important à retenir de ce paragraphe à l'aide de la modélisation thématique. Le résultat peut *ressembler* à la représentation donnée dans la Figure 4-13. Nous avons utilisé le terme "*ressembler*" pour insister sur le fait que la modélisation thématique est une interprétation, et comme toute interprétation, elle peut différer d'un interpréteur à l'autre..

Un liquide oui, mais lequel ? Il est évident que de nombreuses formations sur Mars s'expliquent par l'écoulement d'un liquide, mais lequel ? Un flot massif d'hydrocarbures a d'abord été envisagé pour expliquer les formes d'érosion, mais les résultats des analyses effectuées par les sondes Viking ont rendu cette hypothèse obsolète. L'écoulement de laves très fluides a alors été avancé. C'est une hypothèse valable, mais elle ne peut pas tout expliquer. Mais c'est surtout l'intervention d'une grande quantité d'eau souterraine que l'on privilégie maintenant pour expliquer les formes observées. On peut distinguer deux types de modèles dus à un écoulement d'eau liquide : les chenaux d'inondations et les réseaux de vallées.

Figure 4-12 – Un segment d'un texte scientifique¹

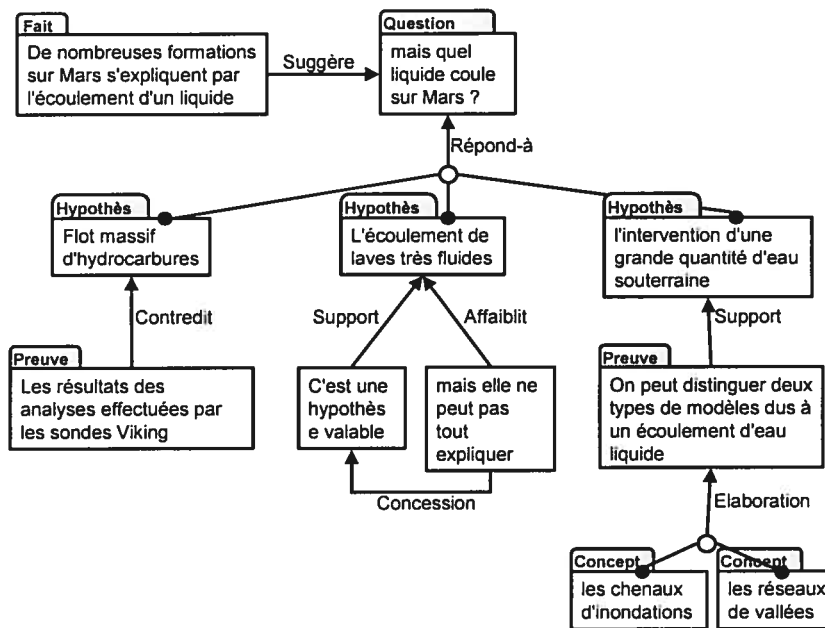


Figure 4-13 – Exemple simple d'analyse thématique

4.3.4 La dimension structurelle

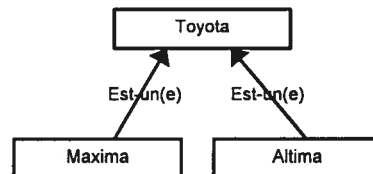
La dimension structurelle tente de capter les aspects structurels des éléments présentés dans le discours. Elle introduit dans la définition des formes épistémiques un ensemble de relations et de constructions structurelles. Cette dimension est similaire à l'analyse

¹ Disponible en ligne sur <http://www.nirgal.net/water.html>

structurelle dans les jeux épistémiques proposé par Collins et Ferguson [Collins et Ferguson, 1993].

Les relations structurelles sont :

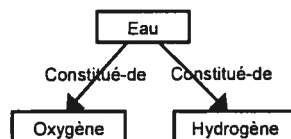
- **Est-un(e)** : c'est une relation ontologique utilisée pour indiquer qu'un élément est d'une certaine classe d'objet. Exemple :



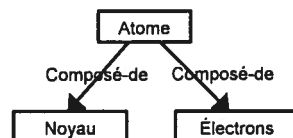
- **Est-membre-de** : c'est une relation d'appartenance à une catégorie ou famille d'objets. Exemple :



- **Constitué-de** : c'est une relation qui permet de décrire les éléments constitutifs d'un autre élément. Exemple :



- **Composé-de** : c'est une relation qui permet de décrire les éléments composant un autre élément. Exemple :

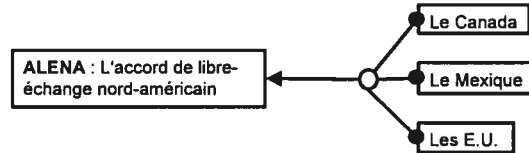


- **Contraste** : c'est une relation qui permet de mettre deux propositions en contraste. Elle permet d'attirer l'attention et la canaliser sur les éléments qui font ce contraste. Exemple:

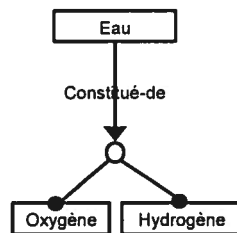


Les constructions structurelles sont :

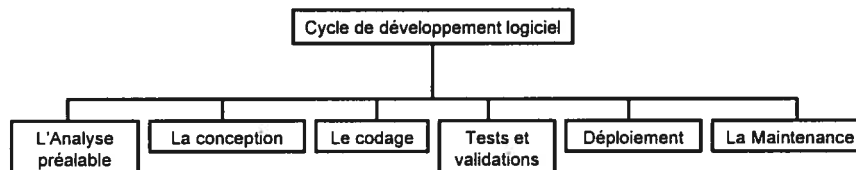
- **La liste** : représente une collection d'éléments rassemblés pour une raison donnée et dont l'ordre n'est pas important. Par exemple:



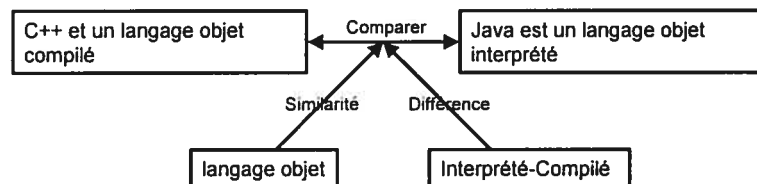
La structure de liste, de même que celle de la conjonction et la disjonction définies par la suite, peuvent être combinées avec des relations binaires 1-1 pour créer des relations 1-n. Reprenons l'exemple utilisé pour illustrer la relation *constitué-de*, il peut être reformulé plus simplement par l'introduction d'une structure de liste :



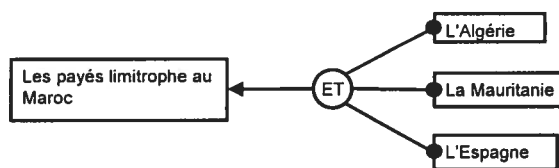
- **La séquence** : représente une séquence d'éléments dont l'ordre a une signification. Généralement cet ordre a une signification temporelle comme pour décrire une suite d'actions dans une procédure ou pour représenter la succession d'événements historiques.



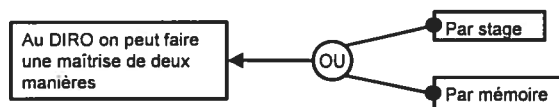
- **Comparaison** : c'est une construction similaire sur le principe à une relation de *contraste* ordinaire, seulement elle permet de rendre explicite les *similarités* et les *différences*. Exemple :



- **Conjonction** : construction similaire à la liste et qui indique la participation conjointe de tous ces éléments. Exemple :



- **Disjonction** : construction similaire à la liste et qui indique la participation disjointe de ces éléments. Exemple :



4.3.5 Modélisation formelle du modèle épistémique

La Figure 4-15 donne la modélisation en notation BNF du modèle épistémique. Dans cette définition le modèle épistémique d'un document didactique est un ensemble de formes épistémiques. Une forme épistémique est une structure de graphe. Les nœuds de ce graphe sont des éléments épistémiques qui peuvent être des micropropositions, des macropropositions, des éléments thématiques (les thèmes) ou des constructions structurelles. Les relations dans ce graphe peuvent être des relations rhétoriques, relations thématiques ou des relations structurelles. La Figure 4-15 donne une modélisation UML de cette définition.

| | |
|---------------------------|--|
| <ModèleÉpistémique> | ::= <FormeÉpistémique>* |
| <FormeÉpistémique> | ::= <ÉlémentÉpistémique>* <RelationÉpistémique>* |
| <ÉlémentÉpistémique> | ::= <Microproposition> <Macroproposition> <Structure> <Thème> |
| <Structure> | ::= <Liste> <Séquence> <Comparaison> <Conjonction> <Disjonction> |
| <Thème> | ::= <TypeThématique> <Proposition> |
| <TypeThématique> | ::= Affirmation Conjecture Hypothèse Fait Axiome Concept Méthode Preuve Problème Question Théorie |
| <Proposition> | ::= <Texte> |
| <RelationÉpistémique> | ::= <TypeRelationÉpistémique> <Source> <Destination> |
| <TypeRelationÉpistémique> | ::= <RelationRhétorique> <RelationThématique> <RelationStructurelle> |
| <RelationRhétorique> | ::= Antithèse Concession Arrière-plan Préparation Facilitation Démonstration Justification Motivation Reformulation Résumé But Cause-délibérée Cause-non-délibérée Circonstance Évaluation Interprétation Élaboration Moyen Résultat-délibéré Résultat-non-délibéré Solution Condition Anti-condition Inconditionnel À-moins-que |
| <RelationThématique> | ::= Affaiblit Renforce Alternative-à Supporte Contredit Répond-à Réfère-à Suggère |
| <RelationStructurelle> | ::= Est-un(e) Est-membre-de Constitué-de Composé-de Contraste |
| <Source> | ::= <FormeÉpistémique> <RelationÉpistémique> |
| <Destination> | ::= <FormeÉpistémique> <RelationÉpistémique> |

Figure 4-14 – Définition d'un modèle épistémique en notation BNF

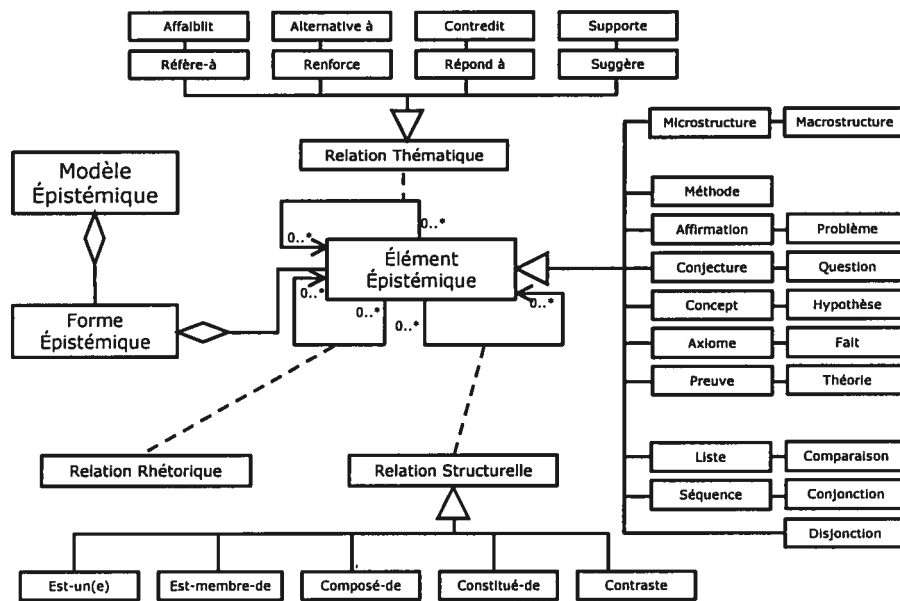


Figure 4-15 – Diagramme de classe UML pour le modèle épistémique

4.4 Le modèle didactique (*LEKC-MI*)

La lecture didactique n'est pas une activité de lecture libre comparée à une lecture récréative. Elle s'insère généralement dans un cadre pédagogique déterminé par le concepteur de cours pour des buts et des objectifs pédagogiques déterminés.

Les théories conventionnelles de l'enseignement, conçues pour l'acquisition de compétences et de performance, ne sont pas convenables dans un contexte de lecture. Un lecteur ne cherche pas à performer, il cherche à *comprendre*. Une des théories d'apprentissage spécialement conçue pour la compréhension est "*The teaching and learning for understanding theory*" ou la TfU, développée à l'université de Harvard par Perkins et Unger [Perkins et Unger, 1999]. Cette théorie a été conçue pour l'enseignement en général, mais à cause de l'accent qu'elle met sur la compréhension comparativement à la performance, nous l'avons adoptée pour son application à l'apprentissage basé sur la lecture didactique.

Si on avait à imaginer une fiction représentant un concepteur de cours préparant un cours basé sur la lecture didactique dans un domaine donné, les épisodes de cette fiction auraient très probablement la séquence suivante :

- La première activité va très probablement consister à identifier les contraintes générales sur ce projet : public cible, niveau de scolaire, niveau de connaissance du domaine, temps alloué pour une telle formation, ... etc.
- La seconde activité va consister à établir qu'est ce qui est intéressant comme sujet à mettre dans cette formation vue les contraintes énumérées ci-dessus. On parle de *sujets génératifs*.
- La troisième activité va consister à définir, pour chacun des sujets génératifs retenus, les notions que les apprenants doivent comprendre lors de cette formation. On parle d'*objectifs de compréhension*. Ces objectifs définissent le "quoi" sans répondre au "comment".
- La quatrième activité va consister à déterminer pour chaque objectif de compréhension, les lectures qui peuvent permettre l'atteinte de ces objectifs de compréhension et les activités d'accompagnement qui peuvent aider et contrôler leur réalisation. On parle d'*activités épistémiques*.

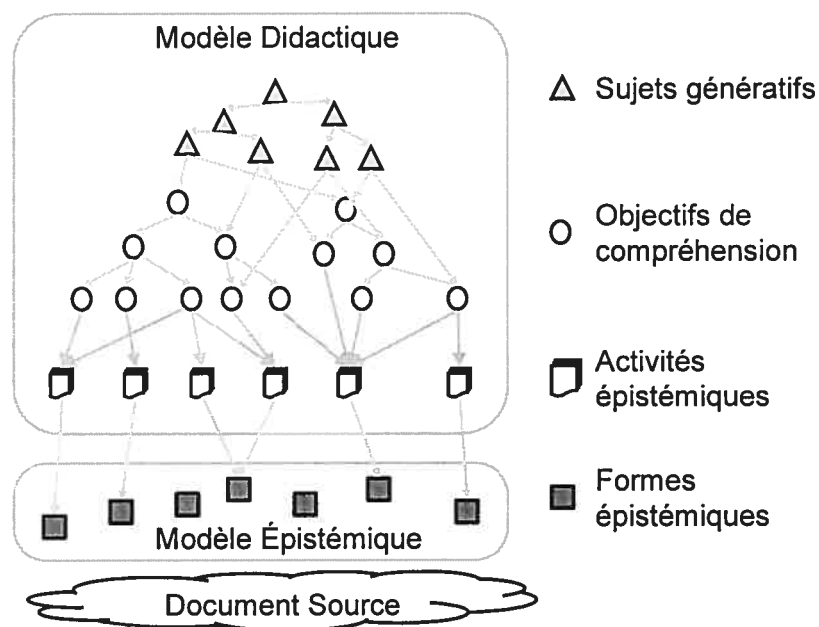


Figure 4-16 – Modèle Didactique

La modélisation didactique dans l'approche *LEKC* permet d'intégrer la dimension didactique dans notre système. Elle permet au concepteur de modéliser la connaissance liée aux sujets génératifs sélectionnés, aux objectifs de compréhension à atteindre et aux activités

épistémiques à réaliser. Le résultat de cette modélisation constitue le modèle didactique *LEKC-MI* (Figure 4-16).

4.4.1 Les sujets génératifs

Dans un processus de formation pour un domaine donné, les sujets qu'on peut traiter peuvent être très variés et diversifiés. Mais à cause des contraintes posées sur de tels projets, surtout celles reliées au temps, une formation ne peut pas embrasser tous les sujets. Une sélection doit donc s'opérer et seuls les sujets importants doivent être retenus. Ce sont les *sujets génératifs*.

Les sujets génératifs sont les sujets importants dans une discipline scientifique donnée, retenus pour les besoins une formation. Leurs critères de sélections peuvent être très nombreux, mais parmi les plus intéressants on a les trois critères suivants à rechercher :

- **Les sujets génératifs sont centraux à leurs disciplines.** Un sujet est central pour une discipline, si l'étude de cette discipline ne peut pas être considérée "complète" sans que ce sujet ne soit abordé. Par exemple si on considère un cours sur l'intelligence artificielle, des sujets centraux peuvent être "*La représentation de connaissance*" ou "*Le raisonnement*".
- **Les sujets génératifs sont accessibles au public cible.** Le choix d'un sujet génératif doit respecter le niveau du public ciblé par la formation. Un sujet peut être accessible *directement* si le public cible possède déjà les connaissances nécessaires pour pouvoir aborder ce sujet. Un sujet peut être accessible *indirectement*, si dans le plan de la formation, le concepteur intègre les sujets préalables (ou pré-requis) qui préparent le public à ce sujet.
- **Les sujets génératifs sont inter-reliés.** Un bon sujet génératif n'est pas isolé des autres sujets dans une formation. Plus un sujet a de connections avec les autres sujets, plus son importance didactique est grande et mérite plus d'attention.

Organisation des sujets génératifs

Les sujets génératifs sont organisés via trois types de relations : les relations de pré-requis, les relations de composition et les relations de référence (Figure 4-17).

Les relations de pré-requis, libellées “*Pré-requis*”, permettent d’indiquer pour un sujet génératif donné S , les sujets génératifs pré-requis qui doivent être présentés au lecteur avant la présentation de S . Cette information est très importante pour la planification pédagogique comme nous allons le démontrer dans les chapitres suivants.

Les relations de composition, libellées “*Composé-de*”, permettent de décomposer un sujet génératif complexe, dit le *super-sujet*, en un certain nombre de *sous-sujets* génératifs de moindre complexité. Un sous-sujet génératif n’est pas réservé à un super-sujet en particulier, mais peut être *réutilisé* dans la composition d’autres super-sujets génératifs. Exemple, le sous-sujet génératif $SS2$ sur la Figure 4-17.

Les relations de références, libellées “*Réfère-à*”, permettent d’indiquer que le contenu d’un sujet génératif donné fait référence au contenu d’un autre sujet génératif. Elles expriment une notion d’interrelation *explicite* entre les deux sujets. Cette notion d’interrelation est parfois *implicite*, comme entre les sujets qui ont en commun des sous-sujets génératifs.

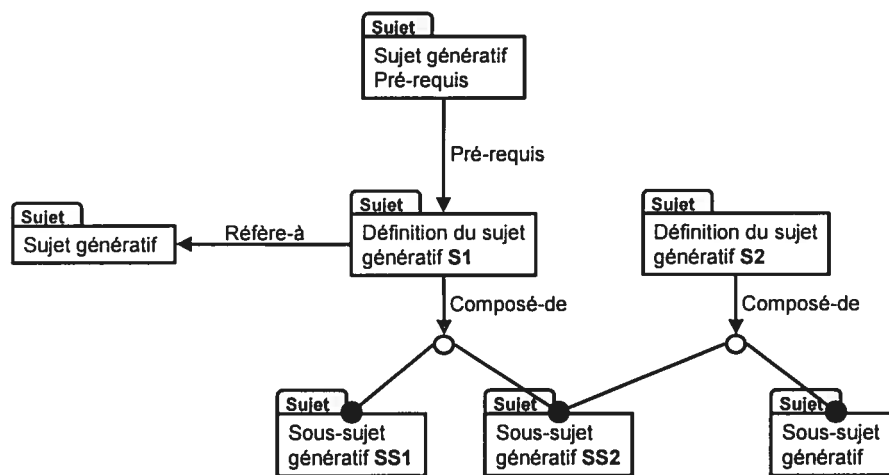


Figure 4-17 – Organisation des sujets génératifs

Les sujets génératifs déterminent le cadre et l’orientation générale d’une formation, mais ils ne donnent pas de détail sur les objectifs à atteindre dans cette formation. Ceci est le rôle des *objectifs de compréhension* associés aux sujets génératifs.

4.4.2 Les objectifs de compréhension

Un objectif de compréhension identifie les concepts, les processus ou les notions que l’apprenant doit comprendre à l’issue d’une formation, relativement à un sujet génératif

donné. Par exemple pour le sujet génératif “*La représentation de connaissance*” en intelligence artificielle, un des objectifs de compréhension pourrait être “*Le lecteur va comprendre la notion de granularité de la représentation*”. Un autre objectif pourrait être “*Le lecteur va apprécier la différence entre représentation intentionnelle et représentation extensionnelle*”.

Les objectifs de compréhension sont des propositions qui expriment le but de compréhension à atteindre sans expliquer le comment. Le format de ces propositions reste libre, mais il est suggéré que ses phrases commencent par des expressions comme “*Le lecteur va comprendre ...*” ou “*Le lecteur va apprécier ...*”. Elles peuvent aussi être exprimées sous la forme de questions ouvertes, comme par exemple : “*Quelles sont les raisons derrière l’occupation de l’Iraq?*”. Ceci contraste clairement avec les définitions conventionnelles pour les objectifs éducatifs qui utilisent des expressions très liées à la notion de performance, comme “*L’apprenant va être capable de ...*”.

Les sujets génératifs et les objectifs de compréhension définissent le “*quoi*” d’une formation respectivement d’une manière générale et d’une manière détaillée. Reste maintenant à déterminer le “*comment*” de cette formation. C’est le rôle des *activités épistémiques*.

Organisation des objectifs de compréhension.

Les objectifs de compréhension sont organisés par l’utilisation des relations suivantes : *Pré-requis*, *Composé-de*, *Alternative-à* et *Sujet-Objectif*.

Les relations *Pré-requis* et *Composé-de* sont similaires aux relations de pré-requis et de composition définies plus haut pour les sujets génératifs.

La relation *Alternative-à* est une relation pour signifier qu’un objectif de compréhension peut être une alternative pour un autre. L’utilité de cette relation est de permettre au système de changer d’objectif quand la réalisation de ce dernier pose trop de difficultés au lecteur.

La relation *Sujet-Objectif* permet d’indiquer l’association entre un objectif de compréhension et un sujet génératif. Un sujet génératif peut être associé à plusieurs objectifs de compréhension, de même qu’un objectif de compréhension peut être associé à plusieurs sujets génératifs.

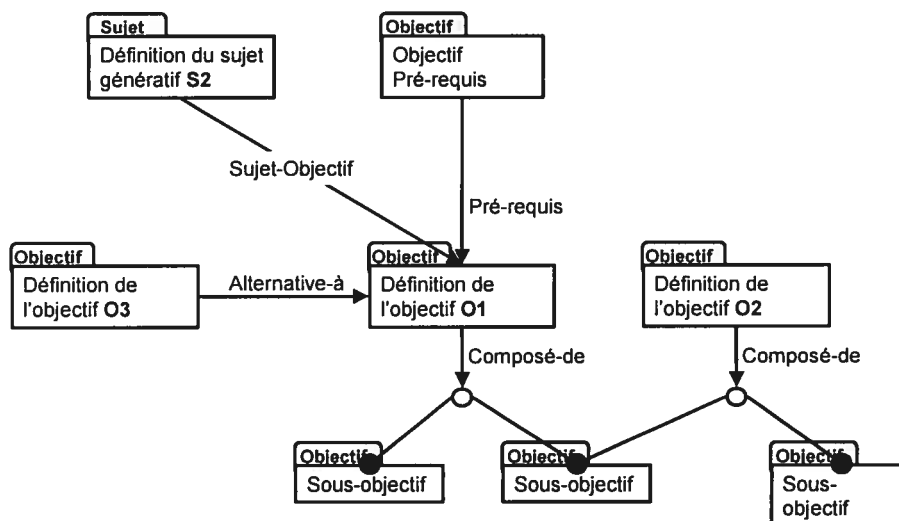


Figure 4-18 – Organisation des objectifs de compréhension

4.4.3 Les activités épistémiques

Rappelons que l'approche *LEKC* est une approche de formation basée sur la lecture de ressources textuelles didactiques. Ses modèles de discours *LEKC-MD* et épistémique *LEKC-ME* représentent l'essentiel du discours véhiculé dans ces ressources sous plusieurs dimensions. Les formes épistémiques, qui sont au sommet de cette modélisation, sont conçues par le concepteur de cours dans le but d'être reconstruites par le lecteur comme preuve de sa compréhension. Une *activité épistémique* définit le cadre de cette tâche de reconstruction.

Une activité épistémique est définie par :

- **Les objectifs de compréhension auxquels elle est associée.** Une activité épistémique peut être associée à plusieurs objectifs de compréhension. De même qu'un objectif de compréhension peut être associé à plusieurs activités épistémiques.
- **La forme épistémique à reconstruire.** C'est la forme épistémique cible.
- **La phrase d'introduction.** C'est un texte expliquant ce qui est attendu de l'apprenant dans cette activité. Il est présenté par le système au début de l'activité et peut prendre la forme interrogative ou la forme impérative.

- **La forme épistémique initiale.** Pour éviter que l'activité épistémique ne soit une distraction pour le lecteur, la tâche de reconstruction n'est pas toujours une tâche de reconstruction complète et intégrale. La reconstruction peut parfois être une réorganisation ou l'achèvement d'une forme incomplète. La forme épistémique initiale est l'état de la forme au début de l'activité. C'est ce qui est déterminé par la méthode de reconstruction.
- **La méthode de reconstruction.** Reconstruire une forme épistémique, qui est une structure de graphe, peut prendre plusieurs allures. Nous avons identifié quatre façons de faire, allons du plus difficile au plus simple :
 - *Construction non assistée* : c'est la façon la plus difficile. Le lecteur n'a au début qu'un indice pour accomplir la tâche, la phrase d'introduction. La forme épistémique initiale est vide.
 - *Construction par organisation* : dans cette méthode, le lecteur reçoit une liste d'éléments qui constituent la forme épistémique à construire (comme une trousse d'outils) et il lui revient de les mettre en ordre. La forme épistémique initiale est vide.
 - *Construction par substitution* : dans cette méthode, le lecteur reçoit une forme épistémique initiale proche de la forme cible mais avec des éléments substitués par des éléments externes (des intrus). Le rôle du lecteur est de les identifier, puis les remplacer par les bons éléments.
 - *Construction par permutation* : c'est la méthode la plus facile. Le lecteur reçoit une forme épistémique initiale contenant exactement les mêmes éléments que la forme cible, mais avec quelques éléments qui ont été permutés. Le rôle du lecteur est d'identifier les éléments permutés et de les remettre en ordre.

Organisation des activités épistémiques

L'organisation des activités épistémiques passe par les relations *Pré-requis*, *Alternative-à*, *Objectif-Activité* et *Cible*.

Les relations *Pré-requis* et *Alternative-à* sont similaires aux mêmes relations définies pour l'organisation des objectifs de compréhension et des sujets génératifs.

La relation *Objectif-Activité* permet d'associer une activité épistémique à un objectif de compréhension. Plusieurs activités épistémiques peuvent être associées à un même objectif de compréhension. Un objectif de compréhension est généralement associé à plusieurs activités épistémiques.

La relation *Cible* permet d'associer une activité épistémique à une forme épistémique cible. C'est la forme à construire dans cette activité.

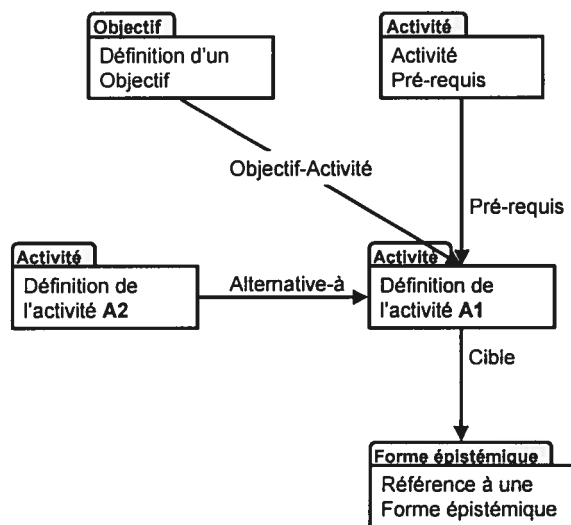


Figure 4-19 – Organisation des activités épistémiques

4.5 Modélisation théorique de *LEKC*

Soit un document \mathcal{T} représentant le texte source sous la forme d'une chaîne de caractères $[c_1 c_2 \dots c_n]$ où $c_{i=1..n}$ est un caractère. Augmenter le document \mathcal{T} c'est lui adjoindre le modèle de connaissance $LEKC(\mathcal{T})$ composé des modèles présentés dans ce chapitre relatifs au discours, à l'épistémique et à la didactique. L'annexe A synthétise la définition théorique de ces modèles.

4.6 Conclusion

L'approche de modélisation *LEKC* que nous avons proposée dans ce chapitre, est destinée à supporter la création de cours basés sur la lecture didactique. À la différence des approches STI traditionnelles dans la modélisation qui étaient orientées principalement vers les

inférences systèmes¹, la modélisation *LEKC* est orientée vers le support actif de l'apprenant dans ses processus cognitifs d'acquisition et de structuration de la connaissance lors de la lecture.

La modélisation *LEKC* considère plusieurs dimensions reliées à la compréhension lors de la lecture, telles les dimensions propositionnelles, rhétoriques, thématiques, structurelles et didactiques. Ce nouveau paradigme de modélisation dans les STIs nécessite une nouvelle organisation des ces systèmes pour exploiter tous les potentiels. Le prochain chapitre exposera une nouvelle architecture de système STI destinée à cet effet.

¹ Effort consenti pour rendre les systèmes capables de raisonnement dans le domaine enseigner

CHAPITRE 5

IT-EKC : UNE ARCHITECTURE DE STI BASÉ SUR L'APPROCHE *LEKC*

Malgré toutes les avancées ces dernières décennies en matière de technologie de l'information, la transmission du savoir, particulièrement dans l'éducation supérieure, continue à se faire via les ressources conventionnelles textuelles, tels les livres de référence et les articles scientifiques, sous format papier ou numérisé (HTML, PDF, présentation PowerPoint, ...etc.)

La numérisation des documents didactiques et leur diffusion électronique sur l'Internet, reste le moyen le plus largement utilisé, pour bénéficier de ces avancées technologiques dans l'éducation. Ceci est dû à plusieurs facteurs, dont les principaux sont les coûts relativement bas de ce genre de projet et la rapidité de mise en œuvre.

L'utilisation des STIs traditionnels, qui se basent principalement sur les approches systèmes experts, n'est pas réaliste dans un contexte où le savoir est en continuelle évolution et augmentation. L'effort, les coûts et l'incertitude sur la viabilité de ce genre de systèmes sont autant d'obstacles qui bloquent leur maturité et entravent leur évolution à un niveau industriel, où ils seraient en mesure de passer le test du *monde réel* préconisé par William Johnson [Polson et Richardson, 1988] :

“If ITSs are to reach their promise, then the laboratory systems must operate and survive in the real world.”

L'approche *LEKC* présentée dans le chapitre précédent adoptait ce principe d'une manière pragmatique, en posant les fondements pour des STIs basés sur l'activité très répandue de la lecture didactique par l'*augmentation* des documents textuels existants.

5.1 La lecture constructiviste

La lecture en tant qu'activité d'apprentissage est un processus très dynamique et qui nécessite la pleine participation de la part du lecteur/apprenant. La théorie constructiviste de l'apprentissage [Bruner, 1973], stipule que l'apprentissage n'est pas un transfert de connaissance *passive* de la tête de l'enseignant à la tête de l'apprenant, mais bien un processus **très actif** où l'apprenant **construit** de nouvelles idées en se basant sur l'état courant et l'état passé de ses connaissances. L'apprenant **sélectionne** l'information, la **transforme**, pose des **hypothèses**, et prend des **décisions**.

Cette définition a été proposée pour l'apprentissage en général, mais elle s'applique parfaitement, à une échelle réduite, au niveau d'une lecture didactique. Le lecteur en lisant ne considère pas toute l'information lue d'un seul coup, il **sélectionne** à chaque fois un sous-ensemble réduit de propositions, **construit** de nouvelles idées en fonction de ses connaissances antérieures, fait des **hypothèses** sur la suite de la lecture ou sur les informations manquantes à ce niveau, et prend des **décisions** sur ces hypothèses à la lecture de nouvelles informations qui peuvent soit confirmer ces hypothèses soit les réfuter.

Une vision analogue pour l'approche constructiviste de l'apprentissage a été proposée par [Ben-Ari, 1998]. Elle stipule que dans un apprentissage passif, comme chaque apprenant apporte une connaissance initiale ou une structuration de cette connaissance différente des autres apprenants, il est très probable que la construction de nouvelles connaissances résultant d'un apprentissage, soit réalisée d'une manière différente d'un apprenant à l'autre. Dans un apprentissage actif, cette construction est constamment supervisée par l'enseignant. Son assistance et ses directives, en plus du feedback des autres apprenants sont autant de moyens pour arriver, ou du moins s'approcher d'une bonne construction de la connaissance. Les mêmes principes peuvent s'appliquer à la lecture. On parle alors de *lecture constructiviste*.

5.2 Connaissance modélisée et son utilisation

La modélisation de connaissance dans les STIs conventionnels, très marquée par l'intelligence artificielle et les systèmes experts, cherchait à décortiquer la matière à enseigner en ses constituants atomiques et élémentaires [Rodenburg, 2001], dans un but bien précis : supporter les inférences du système qu'on voulait à tout prix, rendre "*intelligent*", au détriment d'aspects pratiques et pragmatiques. L'utilisation *directe* par l'apprenant, de cette connaissance modélisée, était un objectif très secondaire. De ce fait la proportion de la connaissance modélisée directement utilisable par l'apprenant était très réduite par rapport à la connaissance modélisée pour servir les inférences du système (Figure 5-1-A). Ceci aurait été valable si ces inférences avaient eu une réelle valeur ajoutée sur les plans pratiques et pédagogiques. Mais l'expérience montre que ce n'est pas le cas et des auteurs commencent à remettre cela en cause [Du Boulay, 2000].

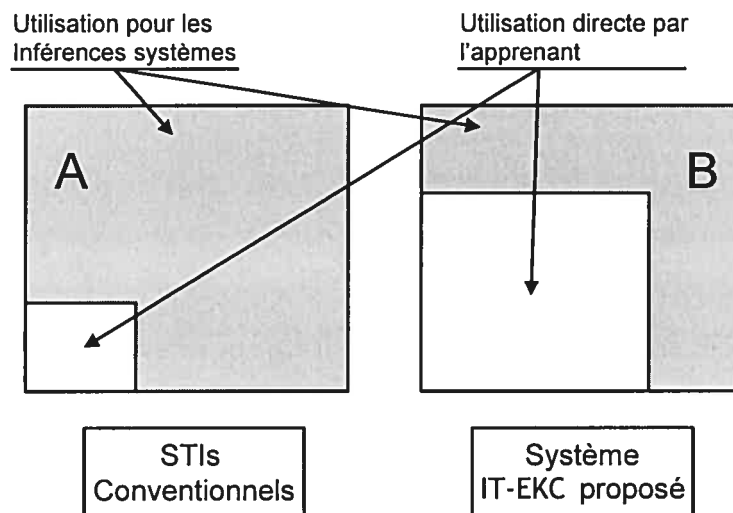


Figure 5-1 – Proportion d'utilisation de la connaissance modélisée

Pour palier à cette situation, nous proposons dans la suite de ce chapitre, une architecture d'un STI basé sur la vision constructiviste de la lecture et sur l'approche de modélisation *LEKC*, et dans lequel l'édition et la modélisation de cours prennent une direction très différente de la démarche classique. Nous l'avons nommé *IT-EKC* pour "*Intelligent Tutoring system based on Explicit Knowledge Construction*".

Selon cette orientation, l'édition n'est plus la production de modélisations servant uniquement aux inférences système, mais la production de ressources d'augmentation qui servent surtout l'apprenant, tout en restant exploitables par le système (Figure 5-1-B et Figure 5-2). Ces ressources d'augmentation sont constituées par les différents modèles du discours et de l'épistémique proposés dans l'approche de modélisation *LEKC* au chapitre précédent.

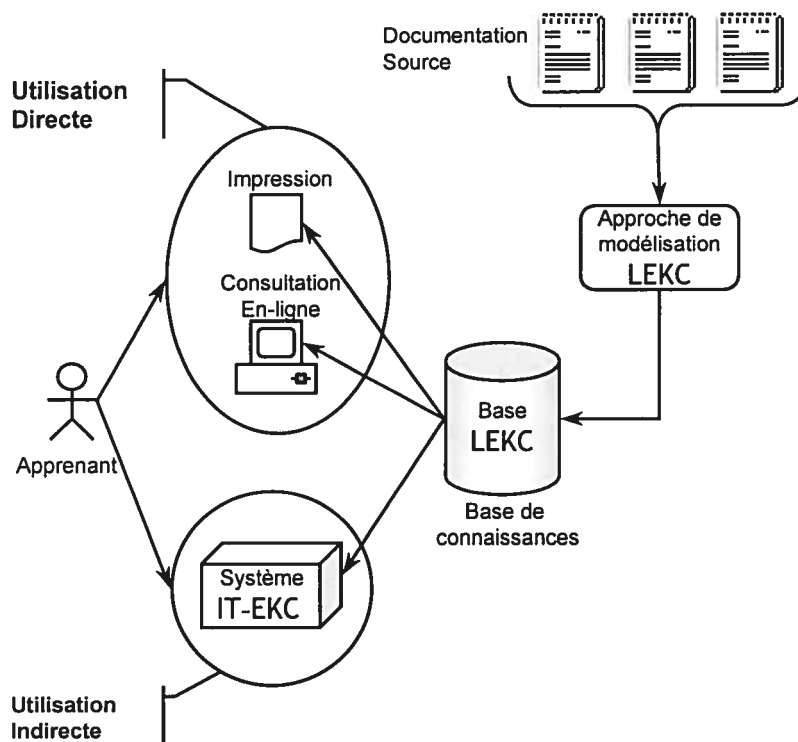


Figure 5-2 – Utilisation de la connaissance modélisée proposée

5.3 Vue globale du système *IT-EKC*

Le système *IT-EKC* proposé, est un STI basé sur l'approche de modélisation *LEKC* et la lecture didactique constructiviste comme moyen de formation. Son architecture peut être décomposée en quatre composantes principales (Figure 5-3):

- **La composante d'édition et de conception**, utilisée par le concepteur de cours pour la modélisation et l'augmentation des documents didactiques.

- La **composante de lecture assistée**, utilisée par l'apprenant (ou le lecteur) dans une session de lecture et qui sert en même temps d'interface d'apprentissage.
- La **composante de collaboration**, est un module pour la gestion d'aspects collaboratifs basés sur la lecture didactique.
- La **base de connaissance LEKC**, est au centre du système, formée par les modélisations de connaissance des trois composantes précédentes.

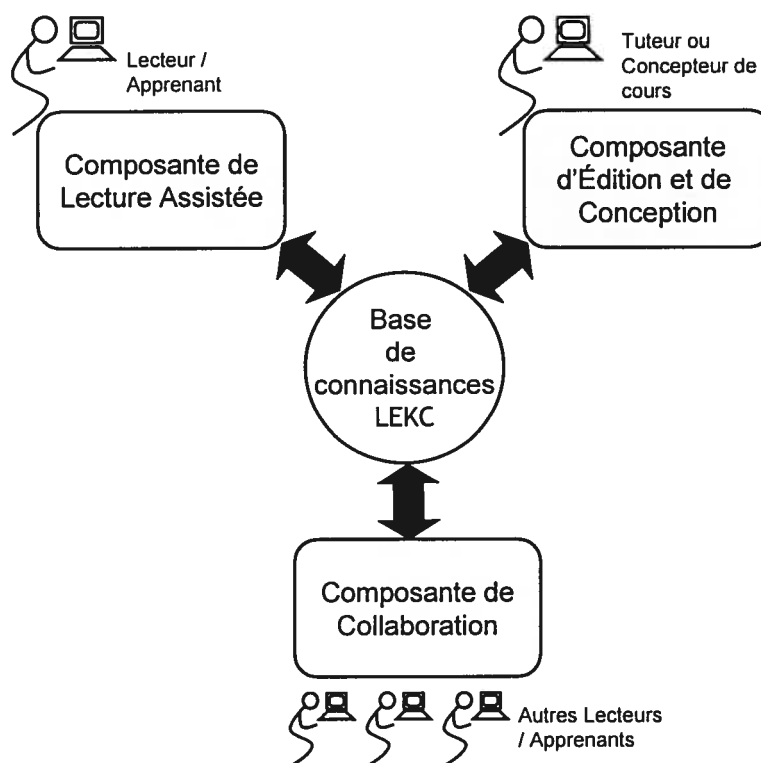


Figure 5-3 – Les grandes composantes du système *IT-EKC*

5.4 La composante d'édition et de conception

La composante d'édition et de conception du système *IT-EKC* est un ensemble d'outils mis à la disposition du concepteur pour la création et la modélisation de cours par l'augmentation des documents didactiques. Le schéma détaillé de cette composante est donné dans la Figure 5-4.

Le travail d'un concepteur utilisant le système *IT-EKC*, commence par la sélection des documents didactiques qui vont faire partie de la formation ciblée. Les documents didactiques sont des ressources HTML en entrée de la composante d'édition et qui sont préalablement éditées par des éditeurs externes¹ de documents HTML, comme Dreamweaver de Macromedia, FrontPage de Microsoft ou par un traitement de texte ordinaire. Les documents ainsi sélectionnés constituent *la base de documents didactiques* du système (Figure 5-4).

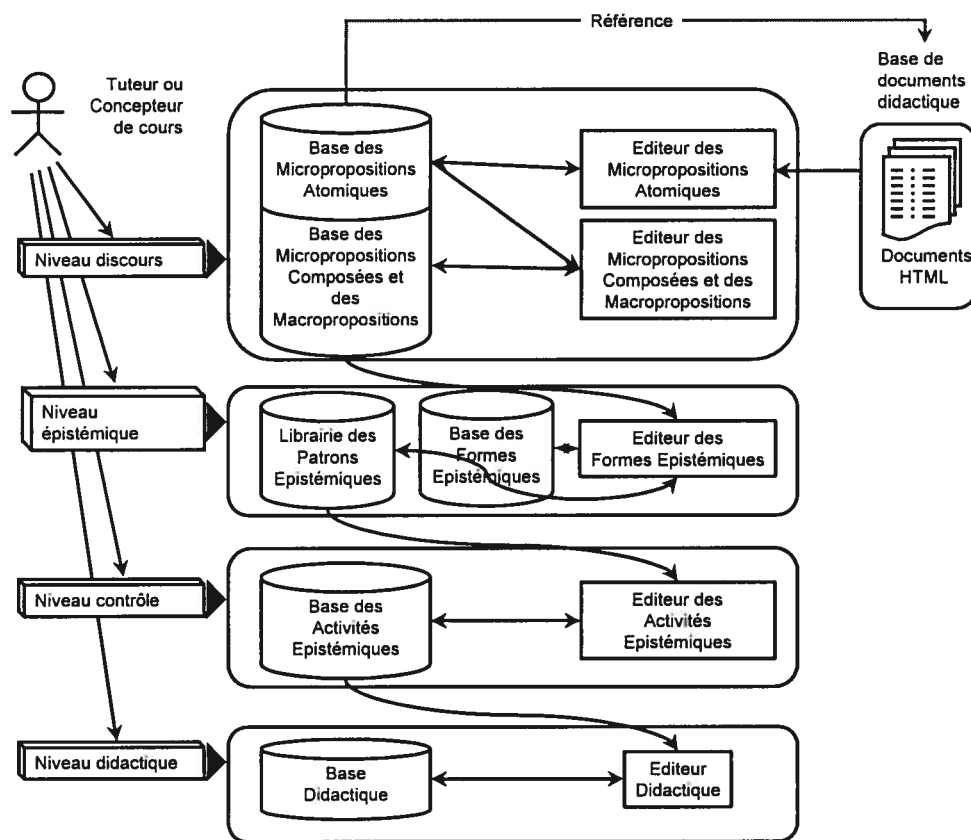


Figure 5-4 – Schéma détaillé de la composante d'édition et de conception

Après cette étape de sélection des documents didactiques, le travail de modélisation (ou d'augmentation de documents) peut commencer. Il se situe à quatre niveaux de modélisation :

¹ Extérieurs à notre application

- la modélisation au niveau du discours,
- la modélisation au niveau épistémique,
- la modélisation au niveau contrôle et
- la modélisation au niveau didactique.

5.4.1 La modélisation au niveau discours

Le niveau discours est relatif à la modélisation des micropropositions et des macropropositions d'un texte didactique. Pour ceci le concepteur dispose de deux éditeurs spécialisés, un pour les micropropositions atomiques et un pour les micropropositions composées et les macropropositions.

Éditeur des micropropositions atomiques

L'éditeur des micropropositions atomiques est un navigateur Web spécial qui permet de sélectionner et de marquer les segments non-décomposables et atomiques d'un texte, qui vont servir aux autres niveaux de modélisation. Le concepteur navigue pour retrouver un document comme il l'aurait fait dans un navigateur Web ordinaire. Une fois le document chargé et affiché, le concepteur commence par sélectionner un segment de texte et demande au système d'en créer une microproposition atomique. Le système crée alors une entrée dans *la base de micropropositions atomiques* en mémorisant quatre informations :

- Un identifiant unique (généralisé automatiquement)
- Le texte du segment
- L'adresse (ou URL) du document source
- L'emplacement du segment dans ce document

Le document source n'est pas modifié et il peut être local comme il peut être n'importe où sur le Web.

Éditeur des micropropositions composées et macropropositions

L'éditeur des micropropositions composées et macropropositions est un éditeur graphique de structures en graphe. Il permet au concepteur de combiner les micropropositions

atomiques marquées sur le document source pour créer les structures plus élaborées que sont les micropropositions composées et les macropropositions. Le résultat de cette édition est sauvegardé dans *la base des micropropositions composées et macropropositions*.

5.4.2 La modélisation au niveau épistémique

Éditeur des formes épistémiques

Le niveau épistémique de la conception est relatif à la création de structures de connaissances sous la forme de formes épistémiques (*représentations externes* en graphe), qui mettent en évidence les notions et les concepts exposés dans le texte. À ce niveau le concepteur dispose d'un éditeur de forme épistémique. C'est un éditeur de graphe spécialisé, offrant une palette de types de nœuds (comme *Concept, Fait, Preuve, ...*) et une palette de types de liens (comme *Est-un(e), Supporte, ...*). Les structures créées, en fonction des types de nœuds et de liens utilisés, permettent la modélisation des dimensions rhétoriques, thématiques et structurelles définies au chapitre précédent.

Formes épistémiques vs. Cartes conceptuelles (*concept maps*)

Les formes épistémiques ne sont pas des cartes conceptuelles (*concept maps*) telles que définies par Novak et Gowin [Novak et Gowin, 1984]. Plusieurs différences existent entre ces deux notions de représentation :

- La nature des nœuds : Les nœuds dans les “*concept maps*” sont uniquement des concepts ce qui est très limitatif. Dans les formes épistémiques, la palette des types de nœuds et des types de liens qui peuvent être utilisés est beaucoup plus riche (*Fait, Preuve, condition, solution...*). Des types de nœuds plus élaborés, comme des animations, des vidéos ou des simulations peuvent aussi être considérés mais à des fins de simplification; nous n'en tiendrons pas compte dans cette version du système.
- La structure générale : La structure d'un “*concept map*” est une hiérarchie où les concepts (nœuds) placés vers le sommet sont considérés comme les plus importants. Nous considérons que la structure de graphe est plus appropriée pour une représentation flexible. L'importance d'un élément dans ce graphe ne découle plus de sa position sur le graphe mais plutôt d'un poids explicitement spécifié par le constructeur.

- Les relations : les relations dans les “*concept maps*” n’existent qu’entre deux nœuds ce qui rend impossible d’exprimer des informations qui portent directement sur une relation. Pour cette raison, nous permettons dans les formes épistémiques que des relations portent directement sur d’autres relations (une relation entre deux relations ou une relation entre un nœud et une relation).

Structures complexes et patrons épistémiques

Les formes épistémiques peuvent parfois être très complexes et difficiles à réaliser. Quand une forme complexe se répète dans plusieurs contextes, le système offre au concepteur le moyen d’en créer un **patron** (ou *template*) qui peut être réutilisé au besoin pour faciliter la conception. Un patron épistémique est une forme épistémique vide, présentant des nœuds et des liens ayant éventuellement un type mais qui ne présentent pas de contenu. Ces patrons peuvent représenter tout un raisonnement, étape par étape, ou la description de tout un processus. Les patrons épistémiques créés par un concepteur sont mémorisés dans la *librairie des patrons épistémiques*. L’utilisation des patrons épistémiques est une façon d’enrichir et d’étendre le vocabulaire de modélisation du système. La Figure 5-5 donne un exemple du patron épistémique pour l’argumentation selon le modèle de Toulmin [Toulmin, 1958].

L’utilisation d’un patron commence par une phase de recherche pour localiser le bon patron dans la librairie des patrons épistémiques. Une fois localisé, le patron est instancié par la création d’une copie identique avec des fentes (*slots*) vides que le concepteur remplira par la suite (voir exemple Figure 5-6).

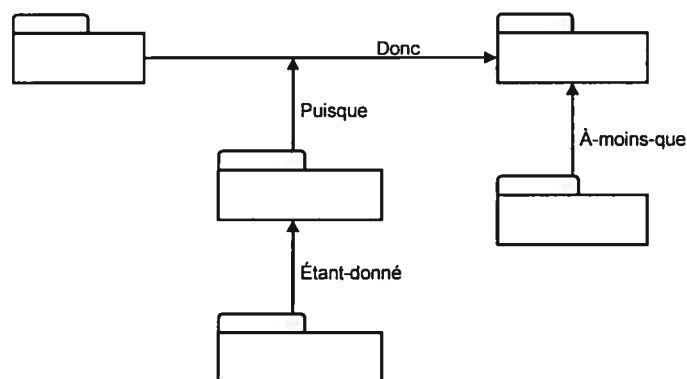


Figure 5-5 – Exemple de patron épistémique pour l'argumentation

Les bibliothèques des patrons épistémiques peuvent mettre en partage le *savoir collectif* de conception par l'utilisation de l'Internet. Des bibliothèques de patrons spécialisées dans une discipline scientifique particulière, accessibles à distance, et contenant les formes épistémiques récurrentes dans cette discipline peuvent être créées et partagées par la communauté de concepteurs de cours dans ce domaine.

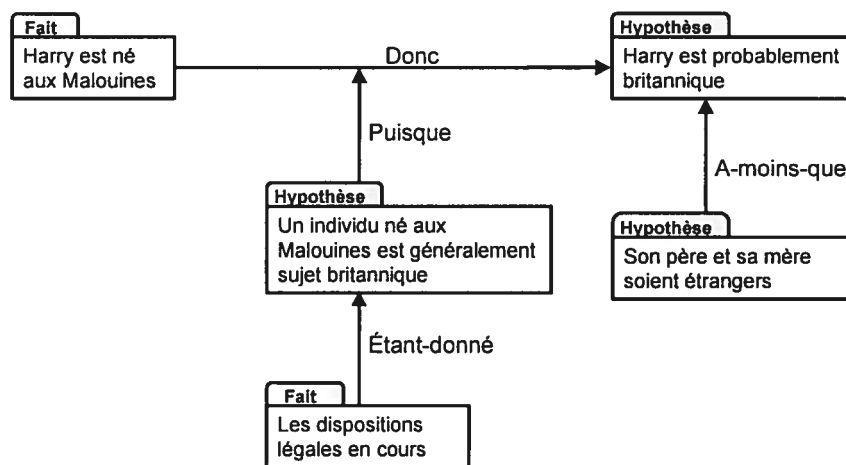


Figure 5-6 – Exemple d'instanciation d'un patron épistémique

5.4.3 La modélisation au niveau contrôle

Les formes épistémiques peuvent être utilisées directement par l'apprenant comme matériel supplémentaire au contenu d'un document didactique. Mais elles peuvent servir aussi à *contrôler* l'activité de lecture et à créer des contextes de réflexion pour le lecteur, si on lui

demandait de les reconstruire ou les compléter. Pour créer ce genre d'exercices de réflexion, le concepteur dispose de l'éditeur d'activités épistémiques. Une instance d'une activité épistémique explique clairement, étape par étape, comment une forme épistémique est reconstruite ou complétée, quel est l'ordre de création préféré, quelles phrases d'aide le système peut présenter à chaque étape, quelle note (ou *score*) apporte une étape réussie dans la note globale, ...etc.

Score et poids épistémiques

Comme nous l'avons définie dans le chapitre précédant, une forme épistémique $\mathcal{F}\mathcal{E}$ est un graphe $\langle \mathcal{N}, \mathcal{A} \rangle$ où \mathcal{N} est l'ensemble des nœuds et où \mathcal{A} est l'ensemble des liens.

Soit $\mathcal{E}_{cor}(\mathcal{F}\mathcal{E}) = \mathcal{N} \cup \mathcal{A}$ l'ensemble des éléments épistémiques de $\mathcal{F}\mathcal{E}$, nœuds et liens confondus. C'est l'ensemble des éléments *corrects* qui font réellement partie de $\mathcal{F}\mathcal{E}$.

Pour signaler les erreurs fréquentes dans la construction de la forme épistémique $\mathcal{F}\mathcal{E}$, le concepteur peut créer une librairie d'erreurs (ou *the bugs library*) qui indique les éléments à ne pas mettre dans cette forme et les raisons de ce choix.

Soit $\mathcal{E}_{err}(\mathcal{F}\mathcal{E})$ la librairie d'erreurs pour la forme épistémique $\mathcal{F}\mathcal{E}$. C'est l'ensemble des éléments épistémiques qui ne font pas partie de $\mathcal{F}\mathcal{E}$ et dont l'insertion dans $\mathcal{F}\mathcal{E}$ est une erreur. Notons que $\mathcal{E}_{err}(\mathcal{F}\mathcal{E}) \cap \mathcal{E}_{cor}(\mathcal{F}\mathcal{E}) = \emptyset$

Soit $\mathcal{E}(\mathcal{F}\mathcal{E}) = \mathcal{E}_{cor}(\mathcal{F}\mathcal{E}) \cup \mathcal{E}_{err}(\mathcal{F}\mathcal{E})$ l'ensemble des éléments corrects et erronés de $\mathcal{F}\mathcal{E}$.

Rappelons qu'une même forme épistémique peut servir dans plusieurs activités épistémiques différentes. Chaque fois, et selon le contexte, le concepteur peut mettre en évidence l'un ou l'autre des éléments de cette forme. Pour prendre en compte cette variation dans l'intérêt que porte le concepteur aux éléments d'une forme épistémique, nous avons prévu deux mesures le $Score()$ et le $Poids()$, qui dépendent de l'activité épistémique $a\mathcal{E}$ en cours.

Soit la fonction $Score : (\mathcal{E}, a\mathcal{E}) \rightarrow [-b..b]$ qui associe à chaque élément de \mathcal{E} un entier entre $-b$ et b . b est une valeur arbitraire fixée par le concepteur. Certains concepteurs préfèrent une notation sur l'intervalle $[-5,5]$, d'autres $[-20,20]$ ou $[-100,100]$. Ce qui est important ici ce n'est pas la valeur absolue de ces scores mais plutôt la valeur relative qu'a un élément par rapport

aux autres éléments de \mathcal{E} . les éléments de $\mathcal{E}_{cor}(\mathcal{F}\acute{E})$ ont des scores positifs et les éléments de $\mathcal{E}_{err}(\mathcal{F}\acute{E})$ ont des scores négatifs.

La fonction $Poids : (\mathcal{E}, a\mathcal{E}) \rightarrow [-100..100]$ est une fonction de normalisation du score. Quelle que soit la valeur de préférence que donne un concepteur à la constante b , la fonction $Poids$ tente de normaliser le score et le ramène à l'intervalle $[-100, 100]$. Elle est définie par le score de l'élément divisé par le total des valeurs absolues de tous les scores.

$$Poids(x, a\mathcal{E}) = \frac{Score(x, a\mathcal{E})}{\sum |Score(i, a\mathcal{E})|, i \in \mathcal{E}} \times 100 \%$$

Notons que :

$$\left[\sum Poids(i, a\mathcal{E}), i \in \mathcal{E}_{cor}(\mathcal{F}\acute{E}) \right] - \left[\sum Poids(j, a\mathcal{E}), j \in \mathcal{E}_{err}(\mathcal{F}\acute{E}) \right] = 100 \%$$

La difficulté épistémique

L'activité épistémique pour construire une forme épistémique donnée peut commencer avec une forme initiale qui contient déjà certains éléments en place. Cela représente une difficulté moindre comparée au cas où la forme initiale est totalement vide. Pour caractériser cette variation de difficulté d'une activité épistémique, nous avons introduit la notion de *la difficulté épistémique*.

Soit $\mathcal{E}_{init}(\mathcal{F}\acute{E}) \subset \mathcal{E}(\mathcal{F}\acute{E})$ l'ensemble des éléments épistémiques initiaux avec lequel l'apprenant commence le travail de construction de la forme $\mathcal{F}\acute{E}$.

La difficulté épistémique d'une activité épistémique $a\mathcal{E}$, notée $Diff_Epi(a\mathcal{E})$, est la différence entre le poids épistémique de la forme finale correcte à atteindre $\mathcal{E}_{cor}(\mathcal{F}\acute{E})$, et le poids épistémique de la forme initiale $\mathcal{E}_{init}(\mathcal{F}\acute{E})$.

$$Diff_Epi(a\mathcal{E}) = Poids_Ep(\mathcal{E}_{cor}(\mathcal{F}\acute{E}), a\mathcal{E}) - Poids_Ep(\mathcal{E}_{init}(\mathcal{F}\acute{E}), a\mathcal{E})$$

$$\text{Avec } Poids_Ep(\mathcal{K}, a\mathcal{E}) = \sum Poids(i, a\mathcal{E}), i \in \mathcal{K}$$

Cohérence de la difficulté épistémique

La difficulté épistémique est maximale, c'est à dire égale à 100%, quand la forme initiale contient toutes les erreurs de $\mathcal{E}_{err}(\mathcal{F}\acute{E})$ et ne contient aucun élément correct de $\mathcal{E}_{cor}(\mathcal{F}\acute{E})$. Dans ce cas $\mathcal{E}_{init}(\mathcal{F}\acute{E}) = \mathcal{E}_{err}(\mathcal{F}\acute{E})$.

$$\begin{aligned}
 \text{Diff_Epi}(aE) &= \text{Poids_Ep}(\mathcal{E}_{cor}(\mathcal{F}\acute{E}), aE) - \text{Poids_Ep}(\mathcal{E}_{init}(\mathcal{F}\acute{E}), aE) \\
 &= \text{Poids_Ep}(\mathcal{E}_{cor}(\mathcal{F}\acute{E}), aE) - \text{Poids_Ep}(\mathcal{E}_{err}(\mathcal{F}\acute{E}), aE) \\
 &= \left[\sum \text{Poids}(i, aE), i \in \mathcal{E}_{cor}(\mathcal{F}\acute{E}) \right] - \left[\sum \text{Poids}(j, aE), j \in \mathcal{E}_{err}(\mathcal{F}\acute{E}) \right] \\
 &= 100\%
 \end{aligned}$$

La difficulté épistémique est nulle quand la forme initiale contient tous les éléments corrects de $\mathcal{E}_{cor}(\mathcal{F}\acute{E})$ et aucun élément erroné de $\mathcal{E}_{err}(\mathcal{F}\acute{E})$. Dans ce cas $\mathcal{E}_{init}(\mathcal{F}\acute{E}) = \mathcal{E}_{cor}(\mathcal{F}\acute{E})$.

$$\begin{aligned}
 \text{Diff_Epi}(aE) &= \text{Poids_Ep}(\mathcal{E}_{cor}(\mathcal{F}\acute{E}), aE) - \text{Poids_Ep}(\mathcal{E}_{init}(\mathcal{F}\acute{E}), aE) \\
 &= \text{Poids_Ep}(\mathcal{E}_{cor}(\mathcal{F}\acute{E}), aE) - \text{Poids_Ep}(\mathcal{E}_{cor}(\mathcal{F}\acute{E}), aE) \\
 &= 0
 \end{aligned}$$

Le poids pour n'importe quelle configuration d'éléments de $\mathcal{F}\acute{E}$, corrects ou erronés, est entre $\text{Poids_Ep}(\mathcal{E}_{err}(\mathcal{F}\acute{E}), aE)$ et $\text{Poids_Ep}(\mathcal{E}_{cor}(\mathcal{F}\acute{E}), aE)$. La relation entre les différents poids est donnée par la Figure 5-7.

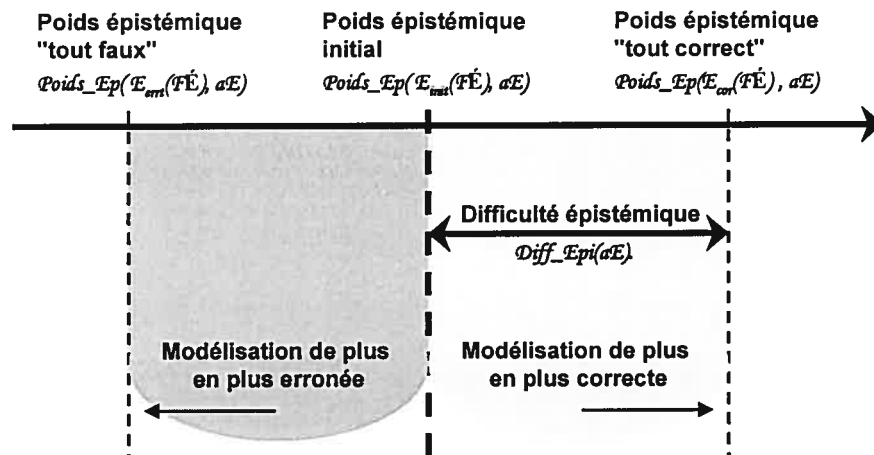


Figure 5-7 – Relation entre les différents poids

Forme épistémique et notion d'explication

La reconstruction d'une forme épistémique par l'apprenant a pour fonction le contrôle et l'orientation de l'opération cognitive de construction de la connaissance. Mais elle a aussi un autre effet, celui de l'explication-à-soi (ou *self-explanation*), effet remarqué dans plusieurs formes de reconstruction de représentations externes par les apprenants [Chi et al., 1989].

[Moore, 1995] dans son travail sur le système d'explication "*Explainable Expert System (EES)*" a identifié quatre éléments pour une bonne explication. Nous les présentons par la suite en faisant l'analogie avec le processus de reconstruction d'une forme épistémique:

- Le quoi (*what*) de l'explication. Dans une activité épistémique ceci est indiqué par la description fournie par le concepteur qui exprime ce qu'il attend de l'apprenant et les objectifs à atteindre.
- Le pourquoi (*why*) ou l'intention de chaque partie de l'explication. Dans notre contexte, le concepteur donne la raison derrière la présence de chaque élément dans une forme épistémique et donne en plus des informations d'aide (les *tips*) que le système peut présenter à l'apprenant sans toutefois lui donner toute la réponse.
- Le comment (*how*) qui exprime la stratégie de présentation. Le concepteur peut spécifier des façons de présentation différentes en variant le contenu de la forme initiale ou l'ordre de création préféré.

- Les suppositions (*what assumptions*) sur les connaissances antérieures de l'apprenant. À l'aide de la notion du pré-requis et du modèle de l'apprenant (que nous allons présenter dans la suite de ce chapitre), le concepteur peut expliciter ses suppositions sur les connaissances que doit avoir l'apprenant avant d'entamer l'activité épistémique en question.

5.4.4 La modélisation au niveau didactique

La modélisation au niveau didactique pour un cours donné est relative à la création, à l'aide de l'éditeur didactique, des sujets génératifs, des objectifs de compréhension et l'association de ces derniers aux activités épistémiques appropriées. C'est la définition du "quoi" et du "comment" d'une formation :

- le "quoi" : est l'ensemble des sujets génératifs et objectifs de compréhension d'une formation,
- le "comment" : est le couplage des objectifs de compréhension et des activités épistémiques.

La séparation du "quoi" et du "comment" au niveau didactique procure deux avantages (Figure 5-8) :

- le "quoi" didactique d'une formation peut se faire indépendamment du reste du processus de conception et probablement par une équipe de concepteurs séparés qui se concentrent sur les orientations stratégiques d'une formation sans être bloqués par les détails. De même, il est possible d'importer la structure du "quoi" d'un organisme ou d'une institution externe, prête à l'emploi.

le même "quoi" peut être mis en contexte dans différents "comment". Imaginons deux livres de références qui peuvent être utilisés pour le même cours. Les "comment" basés sur chacun des livres vont différer, puisque les contenus sont différents, mais le "quoi" reste le même pour les deux. Un apprenant peut choisir l'un ou l'autre pour sa formation, mais le résultat pédagogique reste le même.

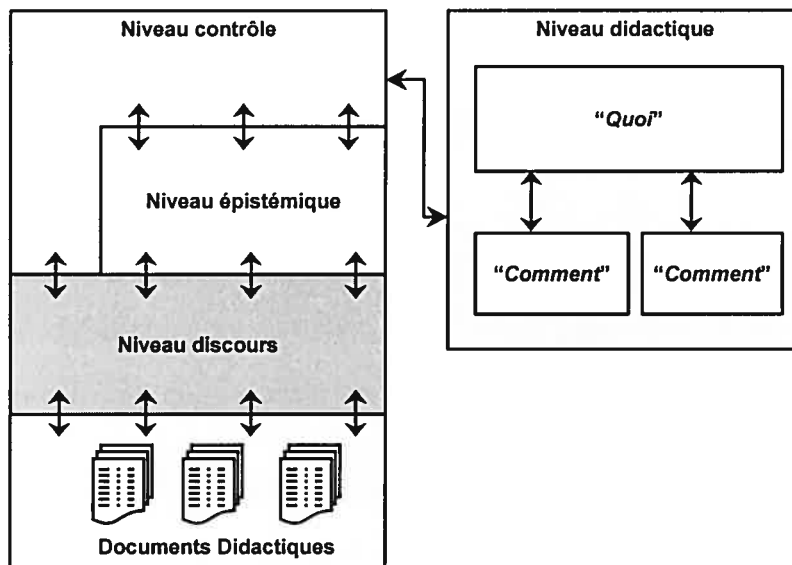


Figure 5-8 – Liens entre les niveaux de modélisation

5.4.5 Diagramme des activités de conception

Pour montrer l'aspect dynamique et fonctionnel de la phase de conception dans le système *IT-EKC*, nous utilisons les diagrammes d'activités UML¹ qui décrivent le comportement et l'acheminement des activités dans un système informatique.

Les principaux éléments d'un diagramme d'activités UML sont les suivants:

| | |
|------------------------------------|--|
| – les activités, | |
| – les transitions entre activités, | |
| – les branchements conditionnels, | |

¹ Unified Modeling Language

| | |
|---|---------|
| – le début, | Début ● |
| – et une ou plusieurs terminaisons possibles. | Fin ○ |

La Figure 5-9 donne le diagramme d'activités de la phase de conception. Le concepteur débute la conception par une activité d'analyse des besoins de formation. C'est une étape cruciale qui se déroule en amont de notre système et dont le but est de déterminer les besoins en formation auxquels la phase de conception doit répondre. Les grandes lignes et orientations à retenir sont consignées dans un cahier des charges. Il n'y pas de méthodes universelles pour conduire cette étape, mais les lignes directrices sont bien décrites dans l'approche de conception des systèmes didactiques *ISD'*.

Une fois l'étape d'analyse des besoins de formation terminée, la conception proprement dite peut commencer. Deux groupes d'activités peuvent se faire en parallèle : les activités didactiques et les activités épistémiques.

¹ Instruction Système Design

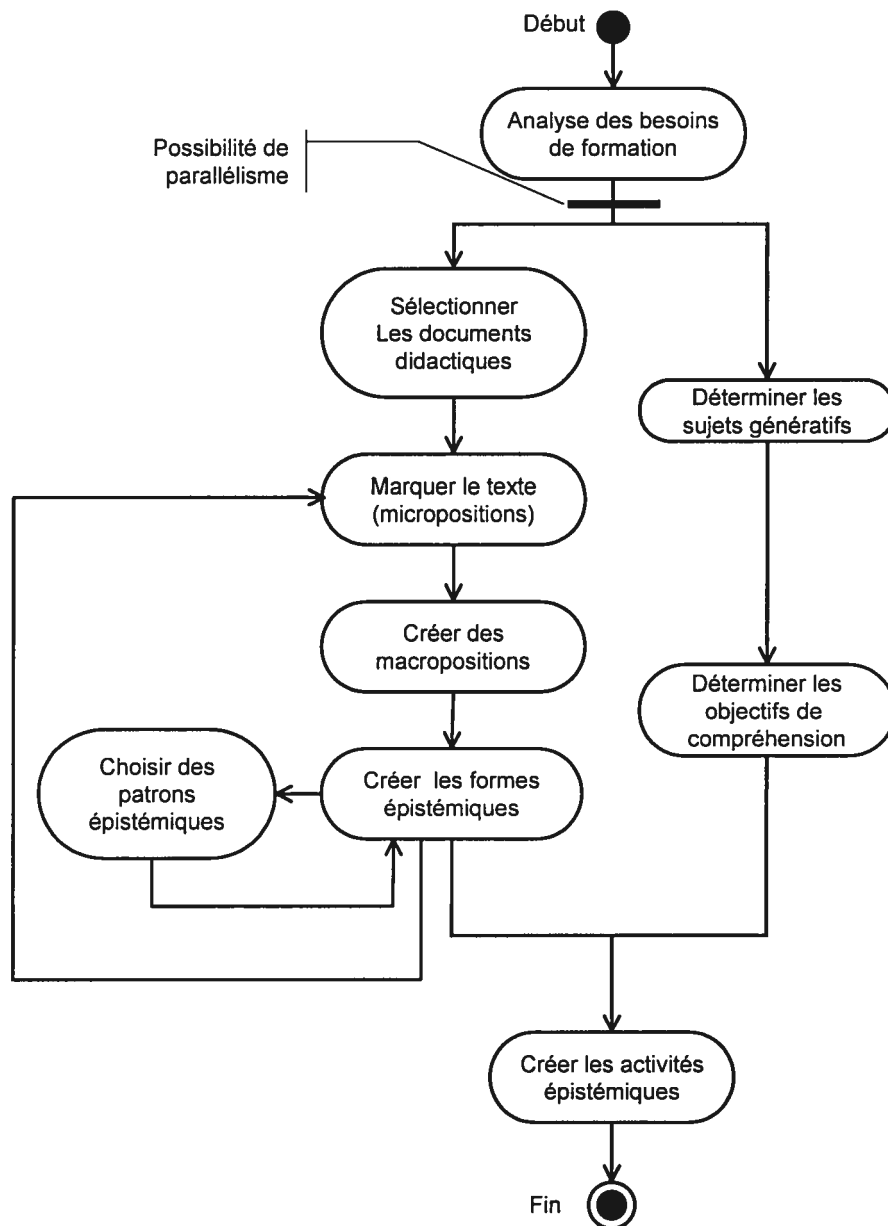


Figure 5-9 – Diagramme d'activités de conception

Dans les activités didactiques, le concepteur doit déterminer les sujets génératifs à traiter par la formation visée ainsi que les objectifs de compréhension pour chacun de ces sujets génératifs.

Dans les activités épistémiques, le concepteur commence par faire une sélection des documents didactiques à utiliser pour la formation et constitue la base des documents didactiques. Des activités de numérisation et de conversion peuvent être nécessaires si ces

documents ne sont pas déjà sous format électronique ou s'ils ont un format autre que HTML. Commence alors des cycles de modélisation, débutant par le marquage des segments de texte qui crée les micropropositions atomiques et aboutissant à la création de formes épistémiques.

Le résultat des activités didactiques et des activités épistémiques est combiné dans l'activité de création des activités épistémiques qui associe les objectifs de compréhension et les formes épistémiques à créer. À la suite de cette étape, les documents didactiques sont prêts à être utilisés par les apprenants/lecteurs grâce à la composante de lecture assistée.

5.5 La composante de lecture assistée

La composante de lecture assistée du système *IT-EKC* est un environnement de lecture utilisé par l'apprenant/lecteur pour suivre un cours basé sur la lecture didactique. Elle permet à l'apprenant de visualiser le matériel didactique, de faire des annotations classiques, comme le marquage et la prise de notes, d'exécuter des activités de réflexion par la construction de formes épistémiques, et de faire la planification de sa lecture.

La composante de lecture assistée est formée de cinq modules :

- le module de visualisation et d'annotation,
- le module de planification,
- le module tuteur,
- le modèle de l'apprenant et
- l'éditeur des formes épistémiques.

La Figure 5-10 donne le schéma détaillé de cette composante.

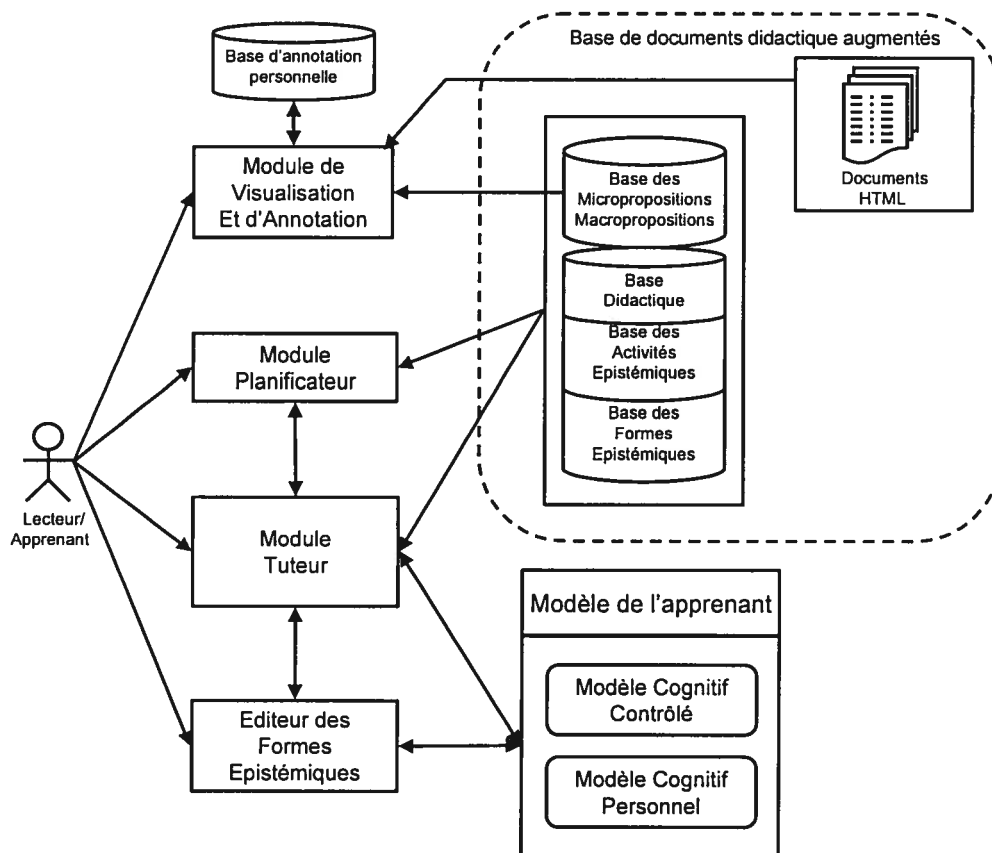


Figure 5-10 – Schéma détaillé de la composante de lecture assistée

5.5.1 Le module de visualisation et d'annotation

Le module de visualisation et d'annotation est une interface semblable à un navigateur WEB qui permet au lecteur de charger, visualiser et naviguer dans un document didactique. En plus de cette fonction d'affichage, le module de visualisation et d'annotation permet de faire les opérations d'annotation classiques, comme le marquage (en jaune dans notre application) des segments de texte qui intéressent le lecteur et la prise de notes textuelles.

Le système utilise l'information du modèle du discours du document en cours d'utilisation (contenue dans la base des micropropositions et des macropropositions) pour faciliter cette opération. Le lecteur visualise les structures graphiques de ce modèle, ce qui lui donne un point de vue supplémentaire sur le sens véhiculé dans le texte. Les macropropositions dans

ce contexte sont comme des notes de lecture, sous forme de petits résumés, avec divers degrés de granularité et prêtes à l'emploi. L'apprenant peut créer ses propres annotations qui vont être stockées dans *la base d'annotation personnelle*.

Une autre aide qu'apporte ce module à l'apprenant est *la visualisation comparative des annotations*. Le système dispose de trois sources d'annotation (marquage et prise de notes): l'annotation faite par le concepteur, l'annotation faite par l'apprenant et l'annotation faite par les autres apprenants. La visualisation des différences entre ces trois sources d'annotation, donne à l'apprenant l'opportunité de réévaluer sa propre annotation.

5.5.2 Le module planificateur

L'utilisation de la composante de lecture assistée du système *IT-EKC* a deux objectifs. Premièrement, s'assurer que l'apprenant a bien lu les documents didactiques qu'on lui a indiqué. Deuxièmement, s'assurer que sa compréhension de cette lecture et semblable, sinon qu'elle s'approche de celle définie par le concepteur du cours dans les différents modèles d'augmentation accompagnant le document didactique. La navigation dans le contenu et la réalisation des activités épistémiques sont les moyens pour atteindre ces objectifs.

Le module de planification a la responsabilité d'organiser et d'ordonner la séquence d'opérations et d'activités que doit effectuer l'apprenant/lecteur lors d'une session de lecture. Généralement, le résultat d'une planification est une suite alternée d'activités de lecture et d'activités épistémiques. Un exemple de cette séquence de planification est donné dans la Figure 5-11.

Nous aborderons en détail dans les chapitres suivants comment le planificateur utilise le modèle didactique et le modèle épistémique ainsi que la connaissance sur l'apprenant contenue dans le modèle de l'apprenant, pour générer un parcours de lecture individualisé et personnalisé pour cet apprenant, en fonction de ses lectures antérieures et des activités épistémiques qu'il a déjà réalisées.

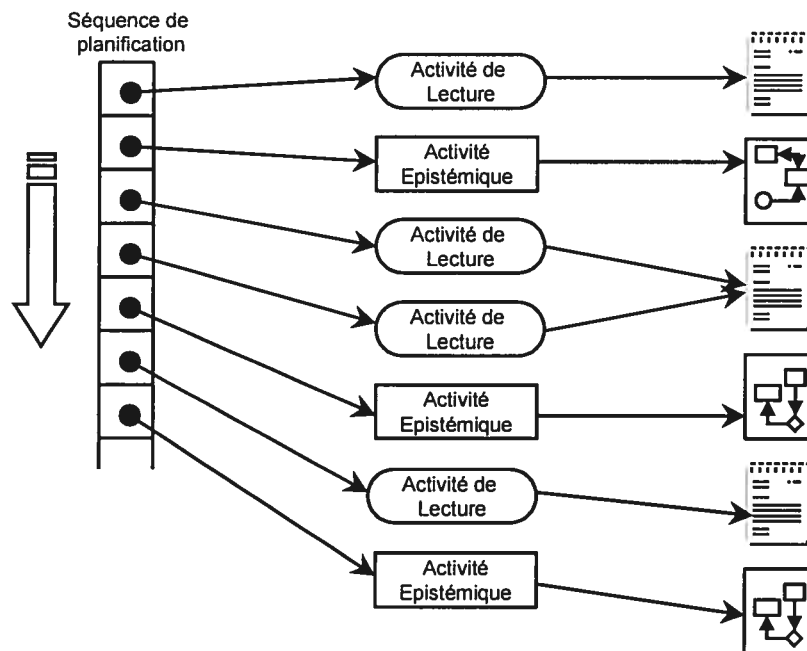


Figure 5-11 – Exemple d'une séquence de planification

5.5.3 Le module tuteur

Le module tuteur est chargé de débiter et de faire le suivi des différentes activités que le système engage avec l'apprenant. En fonction de la planification faite par le module planificateur, le module tuteur sélectionne une à une, les activités à exécuter par l'apprenant. L'ordre d'exécution ne suit pas strictement la séquence de planification générée par le module planificateur. Le tuteur peut prendre l'initiative d'ajouter, de retrancher ou de substituer une activité donnée, pour s'adapter davantage au profil de l'apprenant ou pour faire face à des difficultés dans la réalisation d'une activité particulière (en choisissant une activité équivalente par exemple).

En plus de cette *fonction de supervision*, le module tuteur accomplit aussi la *fonction d'explication*. Il est en mesure de répondre à des questions types que peut poser le lecteur du genre : Que dois-je faire? Pourquoi dois-je le faire? Où suis-je dans la progression? ...etc.

5.5.4 Le modèle de l'apprenant

Pour que le système puisse faire une planification adaptée à l'apprenant et avoir des réactions qui tiennent en compte la différence entre les apprenants (préférences, styles d'apprentissage,...), une modélisation de l'apprenant dans le système est nécessaire.

Le modèle de l'apprenant est la composante du système qui fournit les informations sur l'apprenant. Ces informations sont très variées. Elles peuvent être des informations personnelles, comme le nom de l'apprenant ou son adresse courriel (*email*). Elles peuvent être ses préférences, comme son choix d'interface (le *look-and-feel* de l'application) ou son style d'apprentissage. Elles peuvent être cognitives, comme son niveau de connaissance par rapport au sujet en cours. C'est cette information cognitive qui nous intéresse le plus dans notre modélisation de l'apprenant.

La modélisation de l'information cognitive dans les STIs traditionnels était généralement une information sur la performance de l'apprenant. Elle se faisait principalement par l'attribution de valeurs qualitatives. Par exemple dans le STI *CREAM* [Nkambou, 1996], on utilisait différents vocabulaires pour désigner la performance de l'apprenant, comme : *récite*, *verbalise*, *généralise* et *transfère*. Si cette façon de faire semblait très élégante sur le plan conceptuel, elle posait en revanche de très gros problèmes sur le plan pratique. En effet, on utilisait les résultats des exercices pour déterminer le niveau de compétence ou de maîtrise de l'apprenant. Or un exercice généralement ne concerne pas qu'un seul type de performance¹, mais plusieurs performances et qui interviennent à différents degrés et à différents moments de la réalisation de cet exercice. La note (*le score*) obtenue dans de tels exercices est une note de performance globale qu'on ne peut pas attribuer à un seul type de performance.

La modélisation de l'information cognitive dans le système *IT-EKC* prend une autre approche. Elle se base aussi sur la notion de performance, mais sur ce que nous avons appelé : *la performance épistémique*.

¹ sauf bien sûr pour les exercices très élémentaires

La performance épistémique

La performance épistémique, notée $Perf_Epi(aE)$, est la performance de réalisation ou de construction d'une forme épistémique cible, dans le cadre d'une activité épistémique aE (Figure 5-12).

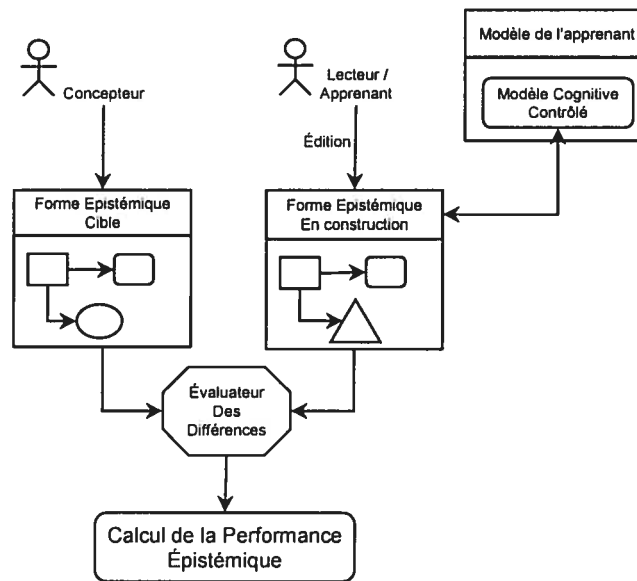


Figure 5-12 – La performance épistémique

Elle peut être mesurée, sans utilisation de valeur qualitative, en comparant la forme épistémique en cours de construction faite par le lecteur/apprenant et la forme épistémique cible à réaliser.

$$Perf_Epi(aE) = \begin{cases} \frac{Poids_Ep(\mathcal{E}_{app}(FÉ), aE) - Poids_Ep(\mathcal{E}_{ini}(FÉ), aE)}{Diff_Epi(aE)} \times 100 & \text{Si } Diff_Epi(aE) \neq 0 \\ 0 & \text{Sinon} \end{cases}$$

Elle est calculée en fonction de la forme épistémique en construction $\mathcal{E}_{app}(FÉ)$, de la forme initiale $\mathcal{E}_{ini}(FÉ)$ et de la difficulté épistémique $Diff_Epi(aE)$. En principe la difficulté épistémique doit être non nulle, car si elle l'est, la forme initiale est alors égale à la forme cible, et dans ce cas l'apprenant n'a rien à construire.

Le modèle cognitif de l'apprenant est composé de deux modèles séparés : un *modèle cognitif contrôlé* et un *modèle cognitif personnel* (Figure 5-13).

Le modèle cognitif contrôlé est un modèle de recouvrement (*overlay model*) qui représente le travail de construction épistémique réalisé par l'apprenant par rapport au modèle épistémique réalisé par le concepteur. Il est contrôlé dans le sens où les formes épistémiques qu'il contient sont reconstruites à la demande du système et correspondent à des formes cibles déjà définies par le concepteur.

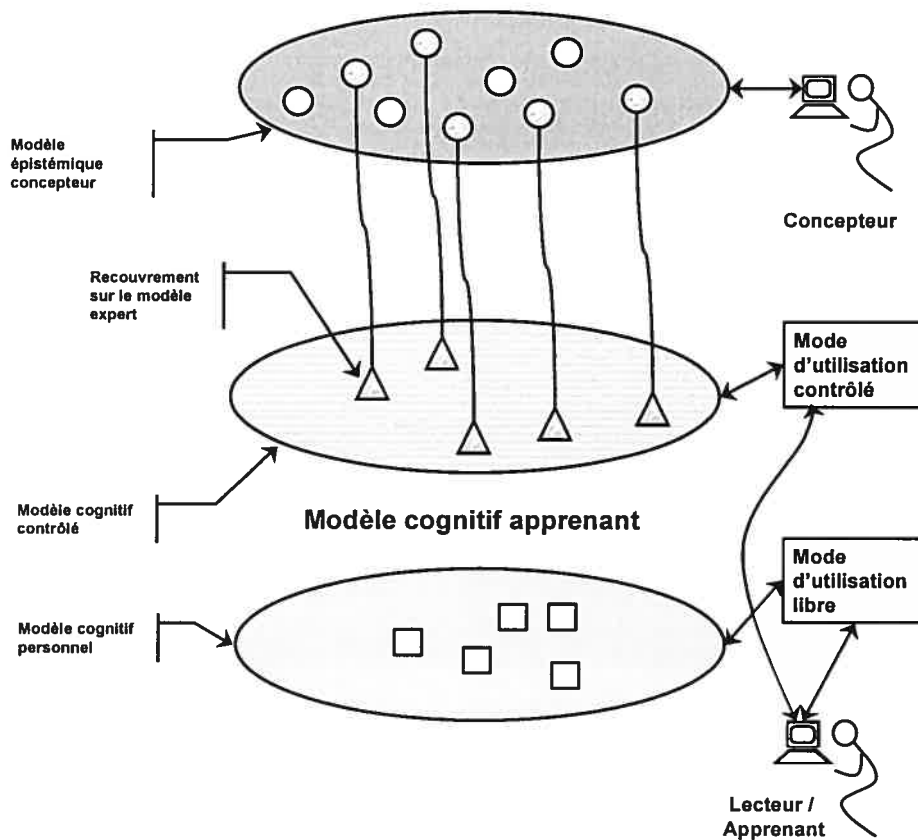


Figure 5-13 – Modèle de cognitif apprenant

Par opposition, le modèle épistémique personnel est un modèle libre où l'apprenant peut exprimer sa propre vision du sujet traité. Mais comme le système ne possède aucune référence de comparaison, l'assistance du système est très limitée dans ce cas. Mais nous verrons comment ce genre de modélisations peut être très utile pour le travail collaboratif et pour l'amélioration continue du contenu et de sa modélisation.

5.5.5 L'éditeur des formes épistémiques

L'éditeur des formes épistémiques utilisé par l'apprenant est très semblable à l'éditeur des formes épistémiques utilisé par le concepteur. La seule différence est que cet éditeur offre deux modes d'utilisation : un mode d'utilisation libre et un mode d'utilisation contrôlée (Figure 5-13).

Dans le mode d'utilisation libre, cet éditeur se comporte exactement comme l'éditeur utilisé par le concepteur. L'apprenant joue le rôle du concepteur et le résultat de cette modélisation constitue le modèle cognitif personnel.

Dans le mode d'utilisation contrôlé, l'éditeur introduit des restrictions pour limiter le travail de l'apprenant à ce qui est demandé par la tâche épistémique en cours d'exécution. Dans les deux modes l'utilisation des patrons épistémiques est possible.

5.5.6 Diagramme des activités de la lecture assistée

L'apprenant peut utiliser l'environnement de lecture assistée de trois façons (Figure 5-14): utilisation directe, utilisation contrôlée et utilisation libre [Rouane, Frasson et Kaltenbach, 2002-b].

Utilisation directe

Au niveau de l'utilisation directe l'apprenant accède directement aux différents modèles d'augmentation qui accompagnent les documents didactiques. Ces modèles de connaissance peuvent être consultés en ligne ou imprimés comme ressources supplémentaires. Le système n'est impliqué dans aucune inférence ou traitement.

Utilisation contrôlée

L'utilisation contrôlée entre dans le cadre de suivi d'un cours préparé par le concepteur ou le responsable de la formation. Elle nécessite une planification pédagogique pour organiser la réalisation des objectifs de compréhension qui sont le but de ce cours.

Comme pour les cours conventionnels, un cours donné via le système *IT-EKC* est réparti sur un ensemble de sessions dont le nombre et la durée peuvent varier selon le rythme et la cadence de chaque apprenant. L'initialisation d'une session de lecture est la première activité

à accomplir, où l'apprenant s'identifie et spécifie quelques paramètres, comme le choix du cours et la durée souhaitée pour la session de lecture qui commence. Vient juste après l'activité de planification initiale. Dans cette activité *le module planificateur* du système dégage une suite d'activités épistémiques à réaliser en fonction des objectifs de compréhension à atteindre, du modèle cognitif contrôlé de l'apprenant et de la durée souhaitée pour cette session de lecture.

La suite des activités épistémiques planifiées est séquentiellement parcourue par *le module tuteur*. À chaque cycle, une activité épistémique est présentée à l'apprenant pour la reconstruction d'une certaine forme épistémique. Si le mode de reconstruction nécessite un guidage pas-à-pas, l'activité épistémique est subdivisée en une suite d'étapes présentées individuellement. À chaque étape, l'apprenant est invité à focaliser sur un seul élément de la forme à réaliser et le contenu approprié pour la réalisation de cette tâche est affiché. Dans les situations où le guidage n'est pas nécessaire, seul le contenu approprié est affiché et l'apprenant est libre de choisir l'ordre de réalisation. Mais dans les deux cas, le résultat de la modélisation faite par l'apprenant, est soumis à une évaluation des différences par rapport au modèle de référence de l'expert, et selon les différences une décision pédagogique est prise par le module tuteur. Les décisions pédagogiques possibles dans le cas où des erreurs seraient détectées dans le travail de l'apprenant sont:

- **Présentation d'indices et d'explications** : en fonction de l'étape en cours, le tuteur présente à l'apprenant des indices ou des explications sans toutefois donner la correction. Les indices et les explications sont préalablement fournis par le concepteur.
- **Lecture ciblée** : en fonction de l'étape en cours le tuteur redirige l'apprenant sur les segments de contenu à relire attentivement pour pouvoir corriger les erreurs constatées.
- **Changement d'étape** : si les deux premières stratégies ne donnent pas les résultats escomptés, le tuteur peut demander un changement d'étape. L'étape courante qui n'est pas terminée sera traitée ultérieurement.
- **Changement d'activité** : si même en changeant les étapes, aucun progrès n'est enregistré, le tuteur peut demander le changement de l'activité. L'activité courante qui n'est pas terminée sera traitée ultérieurement.

- **Demande de replanification** : si malgré toutes ces stratégies, l'apprenant éprouve toujours des difficultés à réaliser les tâches demandées, le tuteur peut demander une replanification où probablement des activités seront substituées et des activités pré-requises insérées.

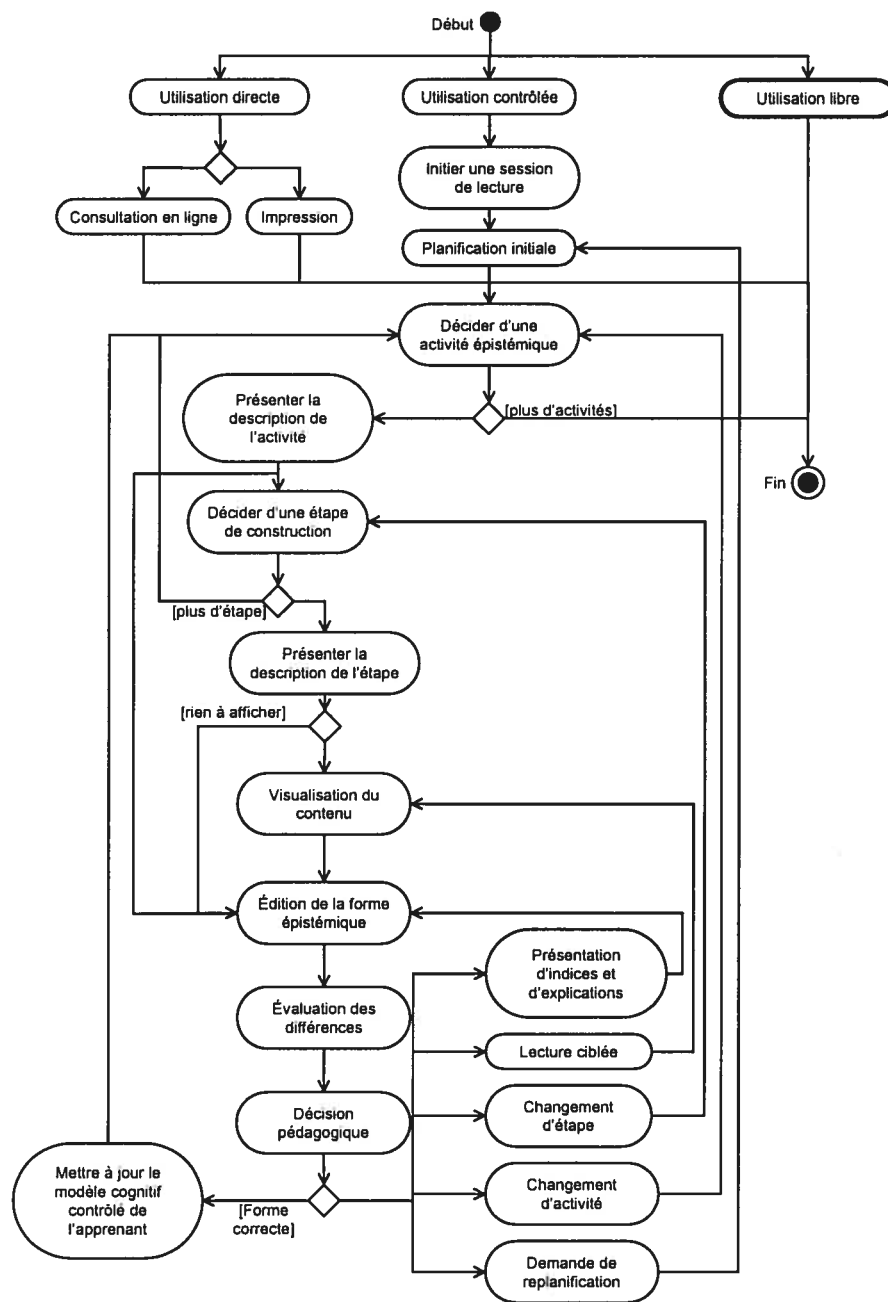


Figure 5-14 - Diagramme des activités de la lecture assistée

Utilisation libre

Dans le cas de l'utilisation libre de l'environnement de lecture assistée, l'apprenant n'est pas limité par les activités planifiées par le système (Figure 5-15). Il a la totale liberté de donner sa propre vision pour le sujet étudié en jouant le rôle du concepteur. Toutefois, la modélisation qu'il fait, se limite aux aspects épistémiques seulement (discours et formes épistémiques) sans toucher aux aspects didactiques (sujets génératifs, objectifs de compréhension, ...).

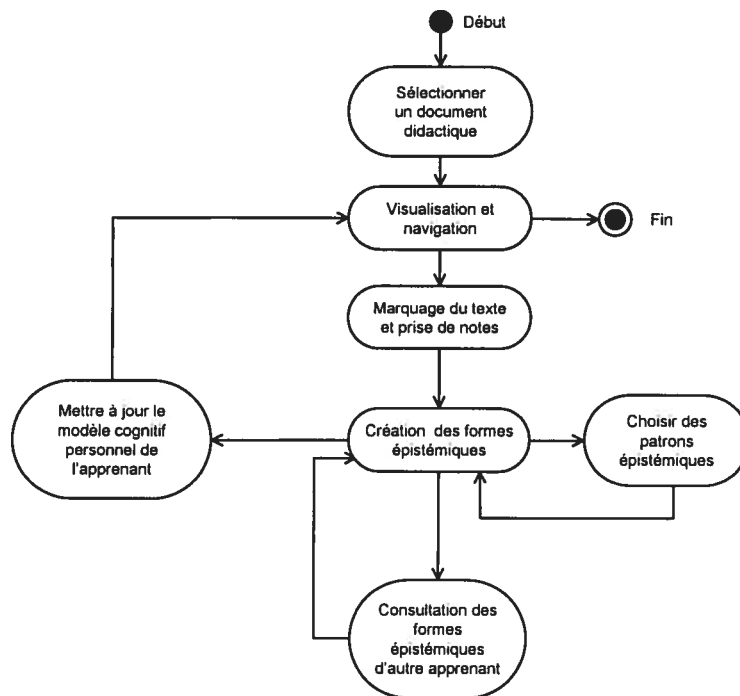


Figure 5-15 - Diagramme des activités utilisation libre

Le travail de modélisation de l'apprenant peut être totalement différent du modèle de référence fait par le concepteur. Dans ces conditions le système ne peut pas analyser les différences comme il l'aurait fait dans le cas de l'utilisation contrôlée, car elles peuvent être très nombreuses. Mais une analyse des ressemblances entre le travail de l'apprenant, le travail de l'expert et les travaux des autres apprenants, donne au système la possibilité de faire quelques actions comme suggérer à l'apprenant de compléter sa forme par les éléments les plus communément utilisés dans le contexte courant donné. C'est un genre d'*auto-completion* pour les formes épistémiques.

5.6 La composante de collaboration

La composante de collaboration du système *IT-EKC* est un ensemble de modules qui permettent aux apprenants de collaborer dans leurs activités didactiques.

Le concepteur dispose de deux modules pour la gestion de la collaboration entre apprenants dans le système :

- Le module de gestion des participants et des groupes, permet de gérer l'information relative aux participants et de définir les groupes auxquels ils appartiennent.
- Le module de définition des activités de collaboration permet au concepteur de spécifier les buts et les objectifs d'un travail collaboratif : quels documents ou séquences de texte les participants doivent lire et quelles activités épistémiques ils doivent réaliser en collaboration.

Pour amorcer une activité de collaboration, le concepteur désigne un groupe de participants choisis dans la base de groupes des participants, et lui associe une activité de collaboration choisie dans la base des activités de collaboration. Deux mécanismes de collaboration sont offerts, un mécanisme de collaboration synchrone et un mécanisme de collaboration asynchrone.

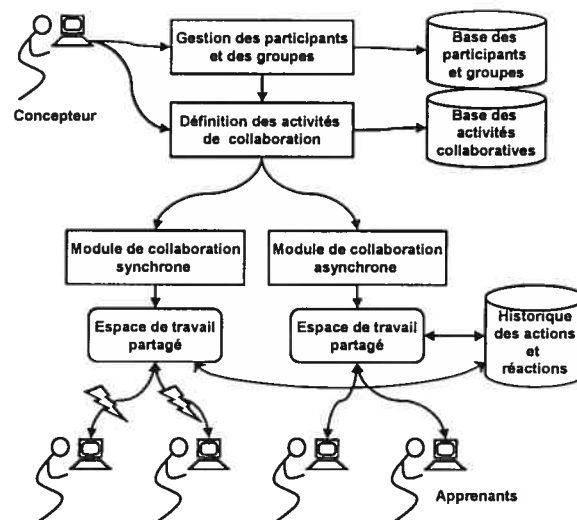


Figure 5-16 – Composante de collaboration

5.6.1 Collaboration synchrone

Dans la collaboration synchrone les participants à un travail de lecture didactique collaboratif sont connectés en même temps au *module de collaboration synchrone* du système. Les interactions et interventions des participants sont directement et instantanément transmises et échangées entre les environnements de travail de chacun (Figure 5-17).

Contrôle de la synchronisation

La synchronisation entre les différents environnements de travail individuels se fait à plusieurs niveaux. Pour gérer cette opération de synchronisation, un mécanisme de contrôle par jeton est introduit. La station de travail qui demande et reçoit le jeton devient *la station maître*. Les autres stations sont dites des *stations clientes*. Les interactions de l'apprenant qui est devant la station maître sont diffusées aux autres stations clientes pour qu'elles se synchronisent avec cette station maître. Les apprenants devant les stations clientes deviennent alors des observateurs des actions de l'apprenant qui a le jeton (Figure 5-18). Pour qu'ils interviennent à leur tour, ils doivent demander et obtenir au préalable le jeton du module de collaboration synchrone. Comme plusieurs demandes au jeton peuvent arriver en même temps, une file d'attente pour ces demandes est utilisée.

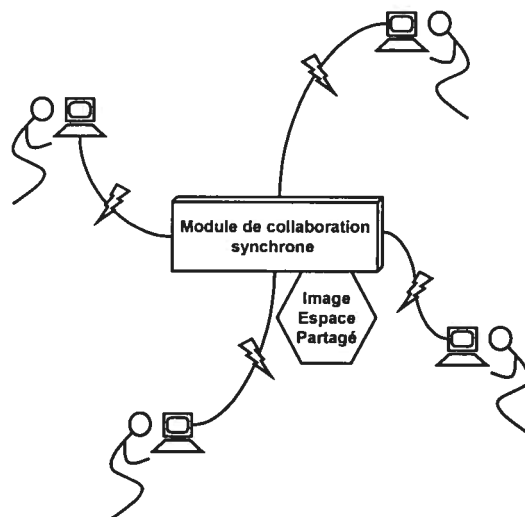


Figure 5-17 – Collaboration synchrone entre apprenants

Types de synchronisations

Chaque participant dispose de deux espaces de travail : un espace privé qui n'est pas visible pour les autres participants et un espace de travail partagé visible des autres participants et qui fait l'objet de synchronisation. Le but de la synchronisation est qu'en tout temps l'espace de travail partagé soit identique sur toutes les stations participantes à cette collaboration. On a trois types de synchronisations :

- Le premier type de synchronisation est la synchronisation de l'affichage et du positionnement. Le module de visualisation dans la station maître diffuse les informations sur le document en cours d'affichage et la position de la portion affichée de ce document. Grâce à ces informations, les autres stations clientes se synchronisent pour afficher exactement la même chose.
- Le deuxième type de synchronisation entre les environnements de travail concerne la synchronisation des annotations, c'est à dire marquages et notes. Quand l'apprenant qui est devant la station maître sélectionne et marque un segment de texte dans le document affiché, l'information est propagée et toutes les autres stations marquent le même segment de texte dans leurs environnements respectifs. Le même principe s'applique pour les notes prises par l'un des participants.
- Le troisième type de synchronisation concerne la synchronisation des actions de construction des formes épistémiques. Généralement, le but du travail collaboratif dans notre contexte, est la réalisation de formes épistémiques compliquées en équipe. Chacun des participants, et à tour de rôle, essaye d'apporter une partie de la solution en apportant des ajouts ou des modifications à la forme épistémique en cours de construction dans l'espace de travail partagé.

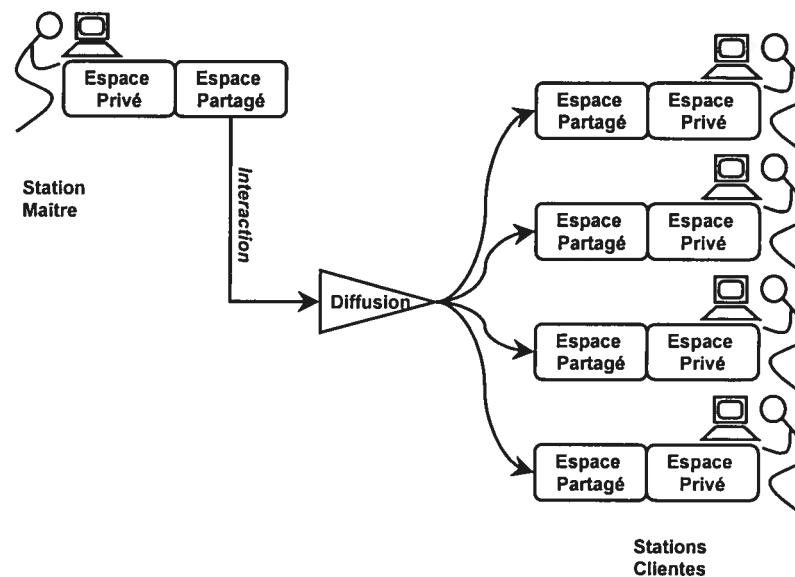


Figure 5-18 – Diffusion des interactions

Historique des interactions

Les actions des différents participants dans un travail collaboratif, ainsi que les réactions qu'elles suscitent sont mémorisées dans la base *Historique des actions et des réactions* de la composante de collaboration. La maintenance de cette base a deux objectifs :

- Premièrement, elle permet de faire des retours en arrière (*le backtracking*) dans le processus de construction d'une forme épistémique compliquée. Après un certain nombre d'actions de construction dans une direction qui semble infructueuse, les participants peuvent décider de revenir sur leurs pas et reprendre la construction dans une autre direction. Cette opération sera expliquée en détail dans le chapitre suivant.
- Deuxièmement, l'analyse du contenu de cette base permet d'enrichir les interventions du système dans le mode d'utilisation non-collaboratif. Dans ce mode où l'apprenant travaille seul devant son ordinateur, le système peut observer les actions de cet apprenant et essayer de les reconnaître dans la base *Historique des actions et des réactions*. Une fois la paire ou les paires action-réaction retrouvées, le système simule une réaction vis-à-vis de l'action de l'apprenant qui prend généralement la forme d'une critique de cette action. Le but est de provoquer et soutenir des contextes de réflexion pour l'apprenant.

5.6.2 Collaboration asynchrone

La collaboration asynchrone est utile pour des participants qui ont des contraintes temporelles qui empêchent leur participation simultanée. Le principe de la collaboration asynchrone est très similaire à celui de la collaboration synchrone. Tous les mécanismes sont identiques sauf pour l'espace de travail partagé et sa synchronisation.

Pour une station participante, il y a deux cas de figures selon que la station est connectée ou non connectée (Figure 5-19).

Quand la station est connectée, et s'il y a d'autres stations participantes connectées au même moment, le système se comporte exactement comme dans une collaboration synchrone pour ces stations.

Quand la station est déconnectée, l'espace de travail partagé n'est plus accessible, mais une image du dernier état de cet espace est toujours disponible localement pour consultation sans possibilité de modification.

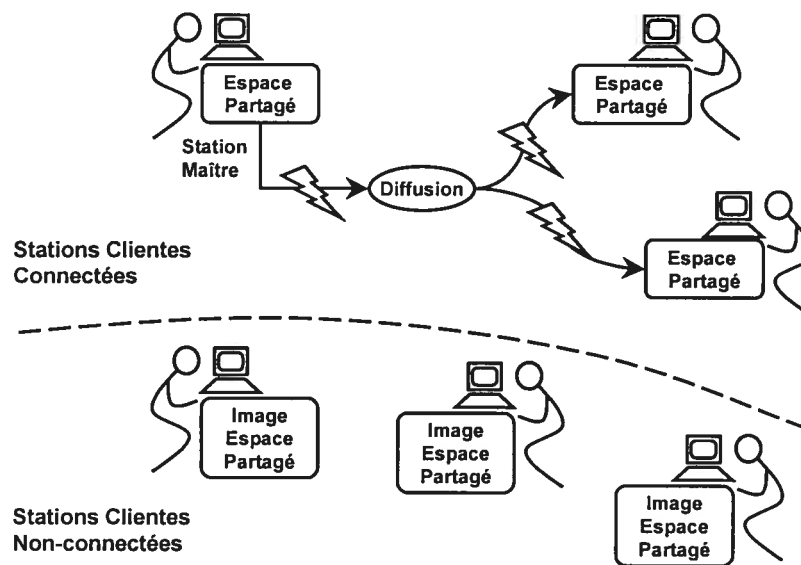


Figure 5-19 – Collaboration asynchrone

5.6.3 Diagramme des activités de collaboration

Les Figure 5-20 et Figure 5-21 donnent les diagrammes des activités de collaboration, respectivement pour la collaboration synchrone et collaboration asynchrone.

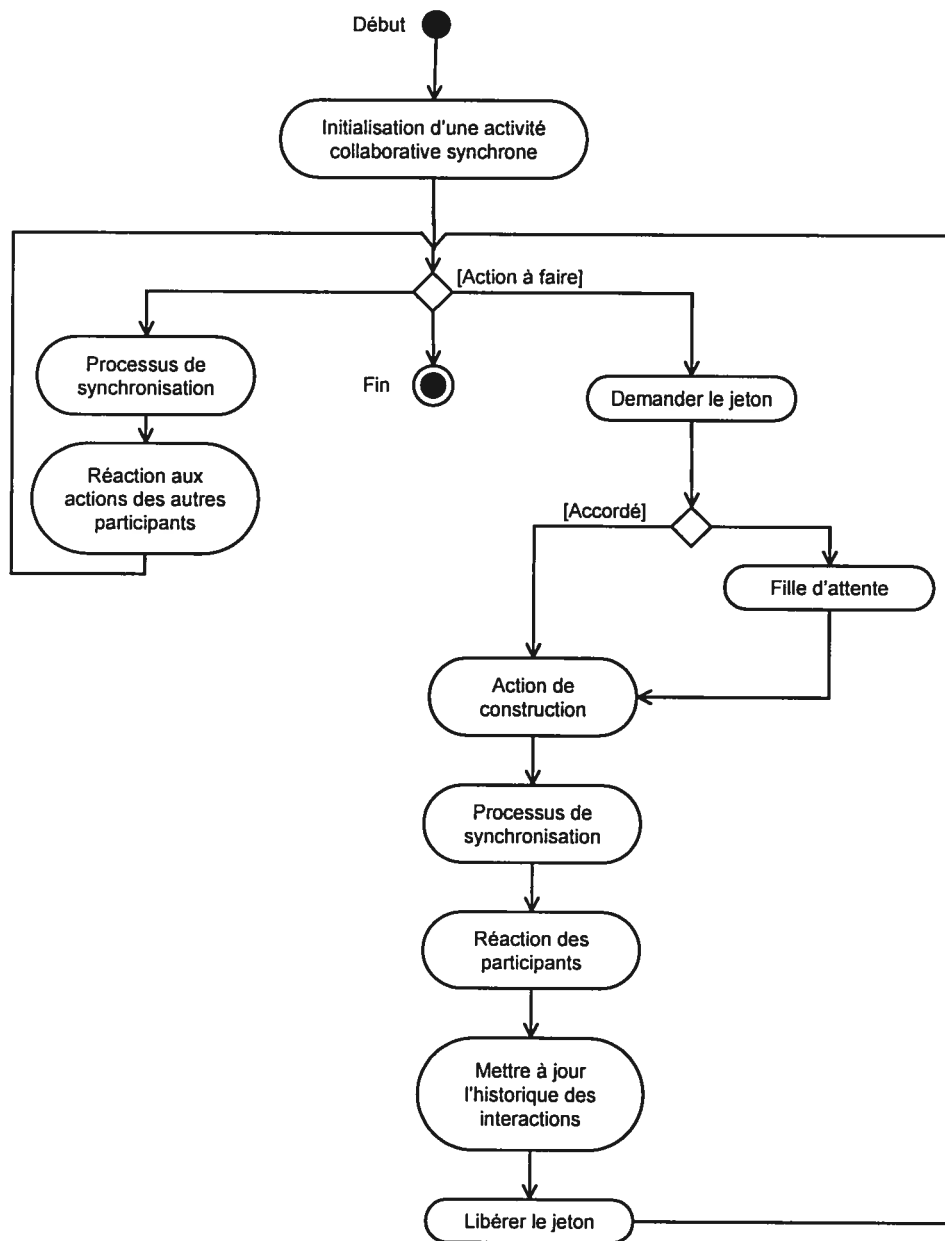


Figure 5-20 – Diagramme des activités de collaboration synchrone

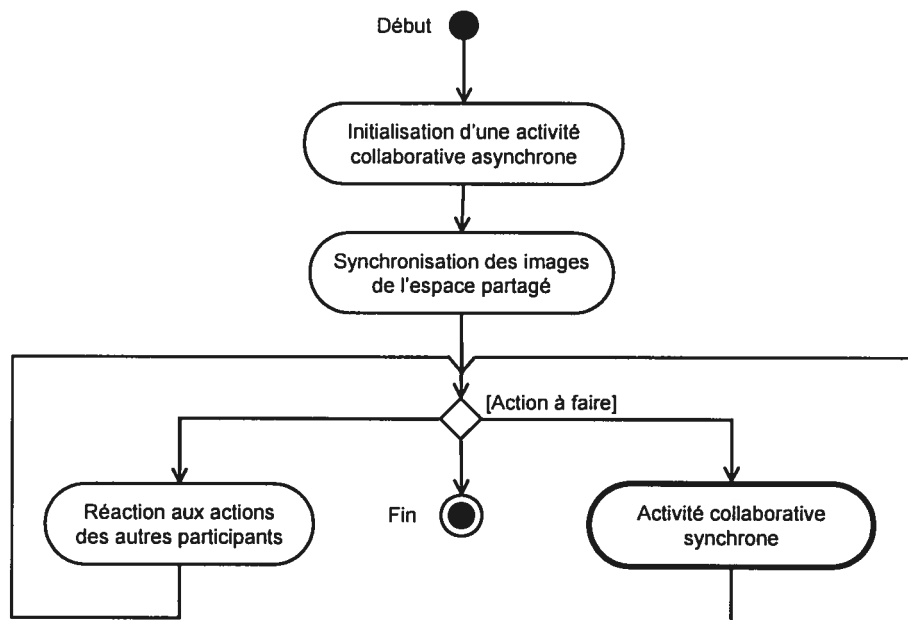


Figure 5-21 – Diagramme des activités de collaboration asynchrone

5.7 Conclusion

L'architecture *IT-EKC* décrite dans ce chapitre montre comment nous pouvons construire un STI complet basé sur la lecture didactique, qui est une activité largement utilisée comme moyen d'acquisition de la connaissance. Chacune de ses composantes a été spécialement pensée et articulée autour des modèles *LEKC* (discours, épistémique et didactique). Dans le prochain chapitre nous explorerons comment cette architecture peut être exploitée effectivement dans la création de cours de même comment elle peut être intégrée dans un STI déjà existant.

CHAPITRE 6 APPRENTISSAGE ET ENSEIGNEMENT BASÉS SUR *IT-EKC*

Dans les chapitres précédents nous avons présenté l'approche de modélisation du curriculum *LEKC* ainsi que l'architecture *IT-EKC* qui donne les grandes lignes pour un STI basé essentiellement sur la lecture didactique et sur la manipulation de représentations ou de formes épistémiques. Cette nouvelle façon de faire nous a obligés à revoir complètement toutes les notions relatives au cours, sa définition et les interactions de l'apprenant avec son contenu. Ce sont ces aspects que va traiter la suite de ce chapitre.

Pour mettre en pratique la majorité des idées énoncées précédemment nous avons mis au point un prototype d'application constitué d'un atelier de construction de cours (modélisation et conception) et d'un atelier de suivi¹ de cours, en utilisant les éléments technologiques suivants:

- Langage de programmation : Microsoft Visual C++ 7 .Net
- Base de données : Microsoft Access
- Dynamic HTML pour la manipulation des pages HTML

¹ "delivery"

Figure 6-1 – Interface de l'atelier de construction de cours

6.1 Exploitation de *IT-EKC* dans la création de cours

6.1.1 La notion de cours dans *IT-EKC*

La notion de cours a de multiples sens et réfère à plusieurs choses à la fois. D'une façon usuelle, un cours est un ensemble de thèmes, de notions ou de sujets qui doivent être abordés dans le cadre d'une formation. Le plan du cours, semblable à une table des matières, regroupe généralement ces éléments d'une façon synthétique.

Un cours dans le cadre d'un STI, est souvent spécifié en terme d'objectifs pédagogiques à atteindre. Ces objectifs peuvent être de deux natures : des objectifs d'enseignement ou des objectifs d'apprentissage.

- Les objectifs d'enseignement sont définis dans les termes et la vision de l'enseignant. Ils décrivent essentiellement les tâches d'enseignement que doit accomplir

l'enseignant. Le plan du cours est une forme implicite d'objectifs d'enseignement qui indique les sujets ou les thèmes que doit aborder l'enseignant lors de son cours.

- Les objectifs d'apprentissage sont exprimés dans les termes et la vision de l'apprenant. Généralement ces objectifs spécifient le résultat attendu d'un apprentissage en fonction des habiletés et des connaissances à acquérir. Les phrases utilisées pour les décrire sont du genre : "À la fin de cette session d'apprentissage l'apprenant sera capable de ..."

Les objectifs pédagogiques à atteindre dans un cours, que se soit d'enseignement ou d'apprentissage, présentent généralement de nombreuses relations entre eux, dont les plus importantes sont les relations de pré-requis et les relations d'ordonnement (par exemple "tel objectif doit être atteint avant tel autre objectif"), [Rouane, 1999]. Ce type de relations a été largement utilisé dans le cadre des STIs pour organiser les objectifs pédagogiques, généralement sous deux formes d'organisation :

- En structures hiérarchiques, comme dans la *hiérarchies d'apprentissage* de Gagné [Gagné et al., 1992], ou
- En structures de graphe, comme dans les *graphes de cours* dans le modèle CREAM [Rouane, 1999] [Nkambou, 1996].

Dans notre système *IT-EKC* nous avons retenu le principe d'organisation en graphe qui utilise le modèle didactique. Comme ce modèle est basé sur la théorie d'apprentissage TfU (ou "*The teaching and learning for understanding theory*", [Perkins et Unger, 1999]) qui est axée sur la compréhension, l'élément essentiel d'organisation d'un cours dans *IT-EKC* est *l'objectif de compréhension*, au lieu d'objectif d'apprentissage ou d'objectif d'enseignement.

6.1.2 L'approche de création de cours dans *IT-EKC*

Un cours dans *IT-EKC* est un "*contrat*" d'enseignement et d'apprentissage basé sur la lecture didactique. Il possède cinq composantes (Figure 6-2): 1) un public ciblé par la formation, qui peut être un apprenant ou un groupe d'apprenants, 2) une spécification des objectifs de formation qui sont les clauses du contrat qui répondent aux besoins en formation, 3) du matériel didactique, essentiellement des documents HTML à utiliser lors des sessions

d'enseignement, 4) Les différents modèles de connaissance *LEKC* relatifs à ces documents, et 5) un tuteur, qui va exécuter et contrôler le déroulement de ce contrat.

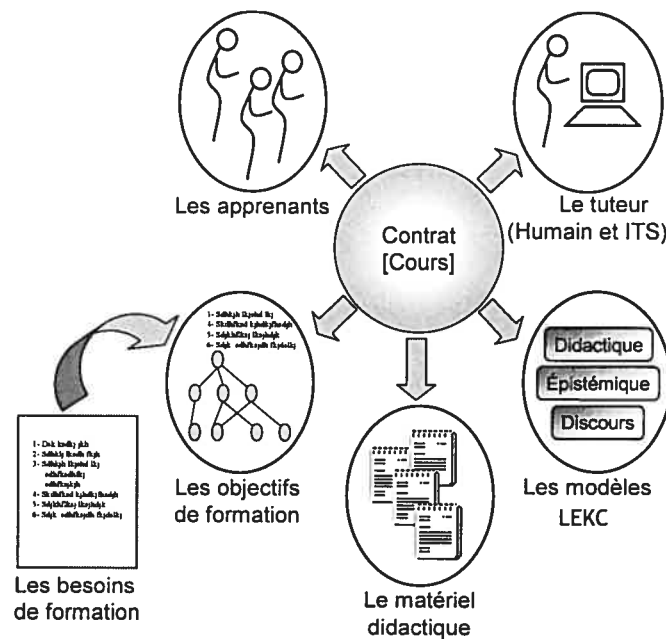


Figure 6-2 – Les composants d'un cours dans *IT-EKC*

6.1.3 Création de la méta-connaissance

La création de la méta-connaissance est le processus de création des différents modèles *LEKC* par le concepteur de cours (discours, épistémique et didactique).

6.1.3.1 Définition des méta-types

La première tâche que doit accomplir le concepteur de cours est la définition des méta-types de la connaissance qui vont servir dans la modélisation. Un ensemble de méta-types est déjà prédéfini dans le système, mais le concepteur est libre d'étendre cette définition en fonction de ses besoins.

Les méta-types que peut créer le concepteur sont de deux natures : des méta-types nœuds et des méta-types liens. Pour les méta-types nœuds le concepteur spécifie (Figure 6-4): 1) une étiquette, 2) une description, 3) la forme géométrique : rectangle, rectangle arrondi ou ellipse, 4) la couleur de fond, 5) les modèles où ce type de nœuds peut être utilisé (discours, épistémique ou didactique).



Figure 6-3 – Exemple de définition de méta-types nœuds

L'exemple ci haut montre l'apparence graphique de la définition du méta-type nœud *concept* et du méta-type nœud *Évidence*.

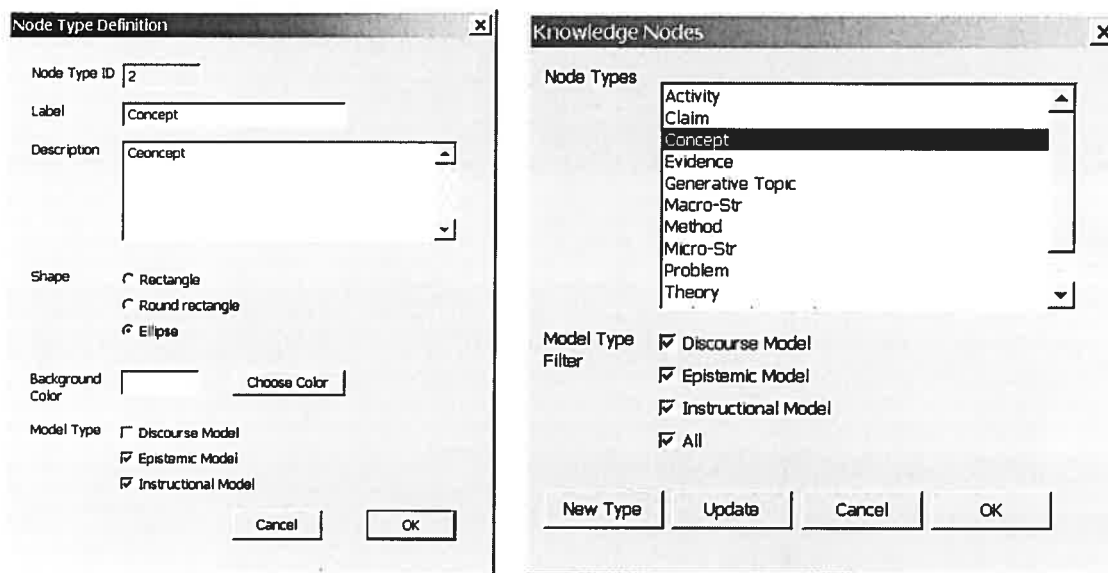


Figure 6-4 – Outils de définition des méta-types nœuds

Pour les méta-types liens le concepteur spécifie (Figure 6-6 et Figure 6-5): 1) une étiquette (le type), 2) une description, 3) un ensemble de qualificatif pour nuancer la force du lien, e.g. *fortement*, *faiblement*, ..., 4) la symétrie, 5) les modèles où ce type de liens peut être utilisé (discours, épistémique ou didactique).

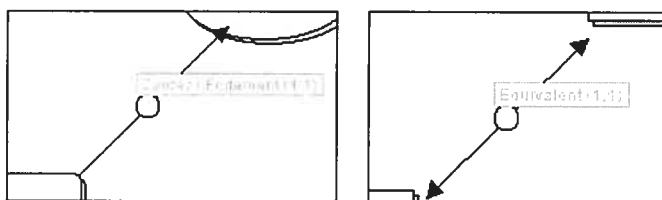


Figure 6-5 – Exemple de définition de méta-types liens

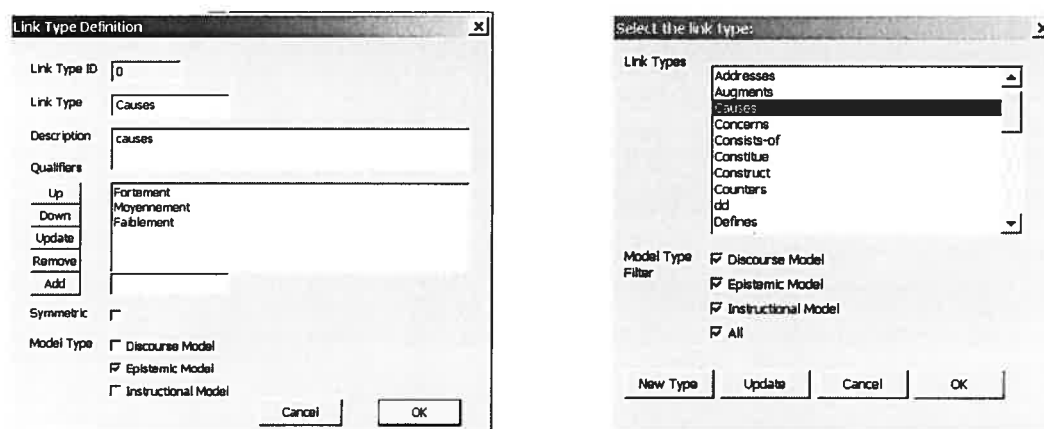


Figure 6-6 – Outils de définition des méta-types liens

L'activité de création des méta-types de la connaissance est généralement faite une fois au début d'un projet de conception et mémorisée dans la base de données. Les définitions faites ne sont pas liées à un cours en particulier mais au domaine de connaissance traité. Elles peuvent servir dans de futures conceptions portant sur le même domaine de connaissance. De même ces définitions peuvent être centralisées par un organisme organisateur (e.g. ACM ou IEEE) et diffusées aux communautés de concepteurs qui travaillent sur des domaines de connaissance similaires. En plus de faciliter le travail des concepteurs, cette harmonisation des méta-définitions est nécessaire pour assurer l'interopérabilité entre les contenus développés par des organismes et des concepteurs différents.

6.1.3.2 Création des méta-tags

Après la définition des méta-types de la connaissance, le travail de modélisation des différents modèles *LEKC* peut commencer. Le premier niveau de cette modélisation est la modélisation du discours. Elle se fait en deux étapes : 1) identification et marquage des séquences de texte sur les documents, c'est la création des méta-tags, et 2) organisation des séquences sélectionnées en structures micro et macropropositions. Nous allons examiner en détail l'étape de création des méta-tags.

La création des méta-tags permet au concepteur de cours d'indiquer au système les séquences de texte à considérer dans le document source. Ces séquences deviennent

marquées visuellement¹ et réactives aux interactions, tels les clics de souris par exemple. Pour assister le concepteur dans la création des méta-tags, nous avons développé un outil spécialisé composé d'un navigateur WEB jumelé à un éditeur de méta-tags (Figure 6-7).

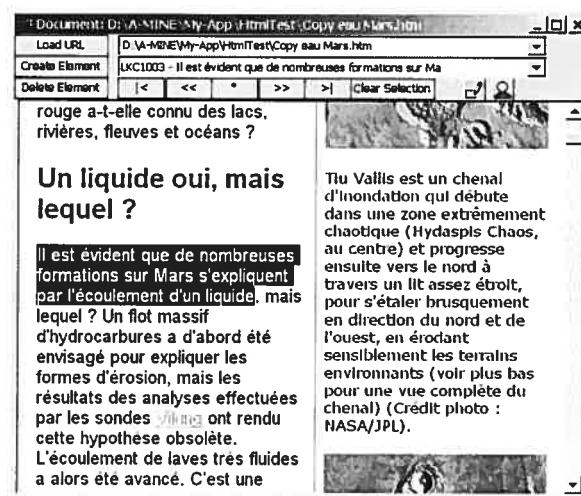


Figure 6-7 – Navigateur et éditeur des méta-tags

Pour créer les méta-tags, le concepteur sélectionne tout simplement la séquence de texte désirée, puis demande au système de générer les méta-tags nécessaires et de les insérer dans la page HTML du document affiché. La Figure 6-8 donne un exemple du code méta-tags généré pour le texte sélectionné sur la Figure 6-7. Cette opération fait associer à cette séquence un identificateur unique (ID) généré automatiquement par le système, et lui crée une entrée dans la base de méta-tags.

```
<span id="LKC1003" style="CURSOR: pointer; BACKGROUND-COLOR: #fefedd">
```

```
    Il est évident que de nombreuses formations sur Mars  
s'expliquent  
    par l'écoulement d'un liquide
```

```
</span>
```

Figure 6-8 – Exemple de code méta-tags HTML

Remarquons que l'insertion des codes méta-tags dans la page HTML du document modifie le document source utilisé et nécessite donc sa sauvegarde. Ceci ne pose pas de problème

¹ en jaune fluorescent dans notre prototype, mais qui peut être changé par l'utilisateur

quand le document est local au système et que le concepteur a le droit de le modifier. Mais par contre on ne peut pas procéder de même pour les documents disponibles en ligne via Internet, car ils ne peuvent pas être modifiés, ce qui interdit leur usage et limiterait sévèrement la portée de notre STI.

Pour résoudre ce problème nous avons introduit un nouveau mécanisme, dit *RDM* pour *Reconstruction Dynamique des Méta-tags* (Figure 6-9), qui permet de marquer des documents sans avoir à les sauvegarder. Son principe est le suivant :

- À la création d'un méta-tag dans l'éditeur des méta-tags, l'entrée associée dans la base des méta-tags mémorise les éléments de la séquence suivants: l'ID de la séquence, le texte de la séquence, l'url¹ du document source, l'ID de l'utilisateur, et l'occurrence de la séquence dans le document (dans le cas où la même séquence se répète plus d'une fois dans le même document).
- Au chargement ultérieur du même document dans notre système, un module spécialisé, le module RDM, utilise la connaissance stockée dans la base des méta-tags pour reconstituer dynamiquement l'ensemble des méta-tags de ce document.

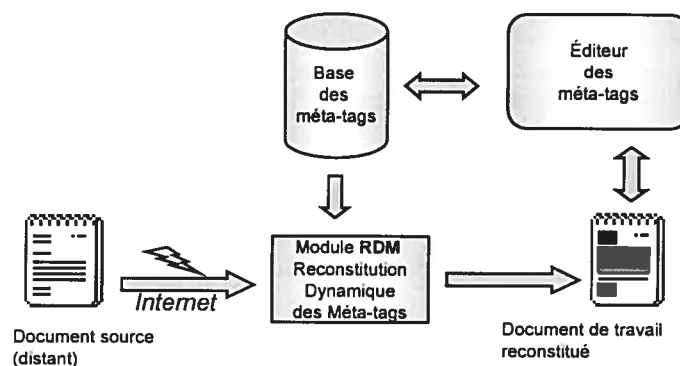


Figure 6-9 – Reconstruction Dynamique des Méta-tags

En plus de permettre l'utilisation des documents distants dans la construction d'un cours sans avoir à les modifier, cette technique permet aussi à plusieurs concepteurs ou même plusieurs étudiants de créer leurs propres ensembles de méta-tags pour le même document,

¹ en fait l'ID du document dans la table des documents qui nous donne par la suite l'url

donnant à chacun la possibilité d'avoir sa propre vision de ce document. L'ID de l'utilisateur mémorisé dans la base des méta-tags est utilisé à cette fin.

Remarque sur l'utilisation de XML.

Le standard XML (*eXtensible Markup Language*) est très utilisé comme langage de balisage extensible. Dans notre contexte, le XML peut offrir beaucoup de robustesse par rapport à l'HTML. Cependant, XML est uniquement un descripteur de données et ne prend pas en considération les aspects de présentation comme dans le HTML.

Dans la réalisation de notre prototype, la décision de travailler avec le format HTML plutôt que le format XML a été prise pour des raisons pratiques, car même si nous avons utilisé XML pour décrire le contenu des documents didactiques, la visualisation de ces documents va toujours nécessiter leur transformation en HTML. D'autre part, beaucoup de composants logiciels (offertes par Microsoft sous format d'Activex) permettent un contrôle très poussé (via le Dynamic HTML) des différents paramètres d'une page HTML ou d'une séquence de texte sur cette page.

6.1.3.3 Création des modèles LEKC

Dans la définition des modèles de connaissance *LEKC* au chapitre 4, nous avons fait une distinction nette et bien délimitée entre les différents types et niveaux de modélisation : discours, épistémique, didactique, rhétorique, thématique, structurel, etc. Cette distinction était nécessaire pour montrer et clarifier le rôle et la nature de la connaissance dans un contexte d'apprentissage basé sur la lecture.

Mais dans la pratique de la modélisation *LEKC*, il est difficile de faire cette distinction, c'est à dire faire une modélisation séparée pour chacune de ces dimensions et les concepteurs trouvent plus facile leur combinaison dans le même plan de travail. Pour cette raison, deux types de modèles seulement sont offerts aux concepteurs : d'un côté *les modèles didactiques* et de l'autre *les modèles du domaine* regroupant les autres modèles (discours, épistémique, rhétorique, thématique, structurel, etc).

L'environnement de modélisation *IT-EKC* offre deux éditeurs graphiques spécialisés respectivement dans l'édition des modèles du domaine (Figure 6-10) et dans l'édition des modèles didactiques (Figure 6-11). Cet environnement offre aux concepteurs de nombreux

outils pour créer et maintenir facilement des structures en graphe représentant des modèles didactiques ou représentant des *formes épistémiques* exprimant les idées véhiculées par les documents.

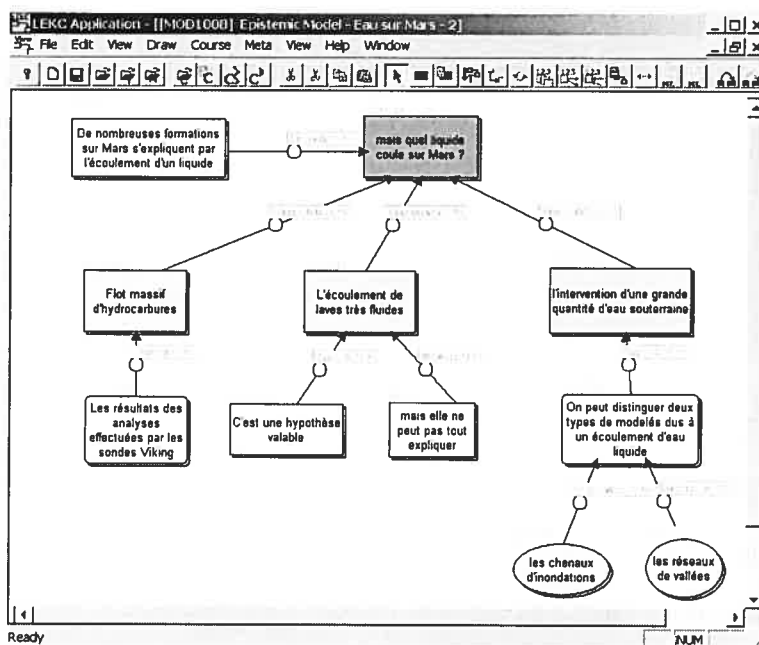


Figure 6-10 – Interface d'édition des modèles de domaine

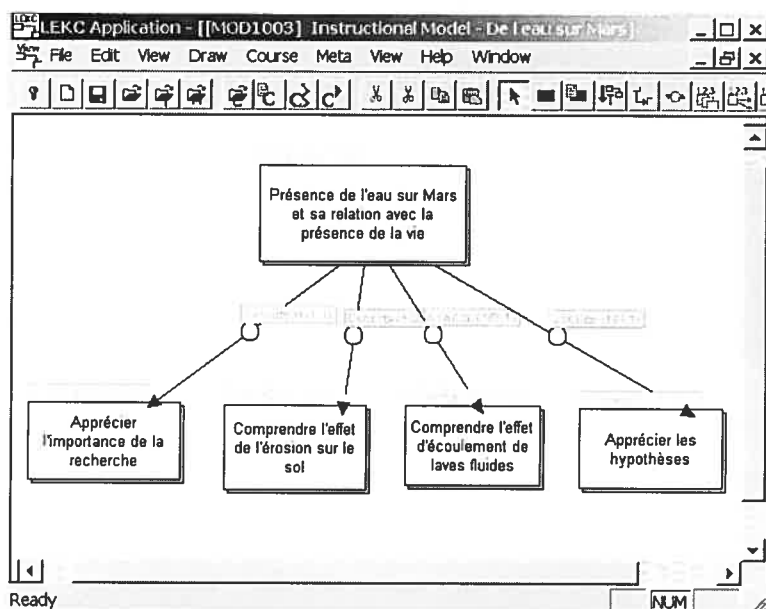


Figure 6-11 – Interface d'édition du modèle didactique

Pour faciliter la visualisation quand le plan de travail¹ devient trop encombré, l'éditeur offre la possibilité d'activer des *filtres épistémiques* pour n'afficher que des éléments d'une dimension donnée (discours, épistémique, rhétorique, thématique, structurel, etc) ou d'un ensemble de types donnés. Un certain nombre de filtres sont déjà prédéfinis et le concepteur peut créer ses propres filtres.

La création des nœuds et des liens dans les formes épistémiques est grandement facilitée par l'utilisation des bibliothèques des méta-types préalablement définies et surtout par l'utilisation des patrons épistémiques prédéfinis qui permettent de créer ces formes par combinaison de parties de structures déjà préparées.

6.1.4 La spécification des besoins et des objectifs de formation

Après la création des modèles *LEKC* pour un document, ce document peut être utilisé comme support didactique pour un cours dans le système *IT-EKC*. La préparation du cours commence par la spécification des besoins en formation. Un cours n'a pas d'intérêt s'il ne répond pas à des besoins réels et identifiables et il est crucial de bien définir ces besoins pour ne pas induire en erreur tout le processus de développement qui vient par la suite.

La spécification des besoins en formation est réalisée sous une forme informelle où ces besoins sont décrits sous forme textuelle. Ils sont revus et corrigés jusqu'à l'approbation par les différentes parties concernées, après quoi ils sont transformés en objectifs de formation qui spécifient avec précision ce qui doit être fait au niveau de la formation pour répondre à ces besoins.

La spécification des objectifs de formation dans le système *IT-EKC* se fait à l'aide des éléments du modèle didactique (Figure 6-12). Le responsable de la formation fait associer à chaque besoin en formation exprimé, des éléments du modèle didactique (sujets génératifs et/ou objectifs de compréhension) qui peuvent répondre à ce besoin.

¹ la fenêtre de travail

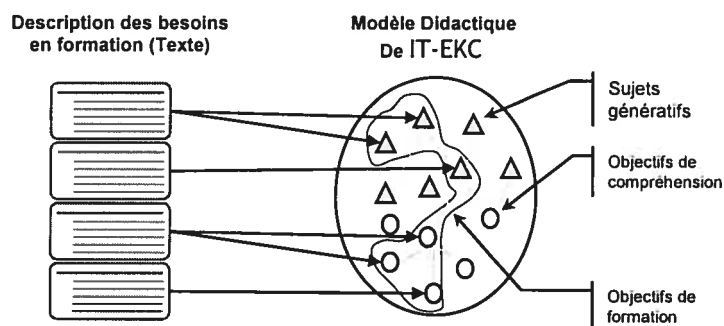


Figure 6-12 – Spécification des besoins de formation

Selon que les documents pédagogiques à utiliser dans la formation sont prédéterminés¹ ou non, le responsable de la formation dispose de deux approches pour spécifier les objectifs de formation : l'approche *Model-based* et l'approche *Document-based*.

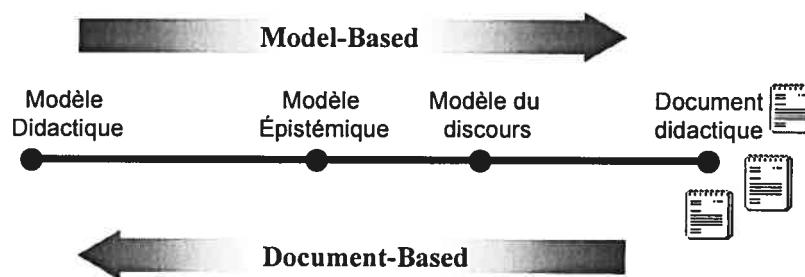


Figure 6-13 – Approches de spécification des objectifs de formation

6.1.5 L'approche Model-based

Dans l'approche *model-based*, les documents pédagogiques à utiliser ne sont pas prédéterminés. Le responsable de la formation réalise la spécification des objectifs de formation à l'aide des éléments du modèle didactique (Figure 6-14), mais sans spécifier à l'avance les documents didactiques à lire. C'est au système que revient la tâche de déterminer ces documents au moment opportun.

¹ Fixés

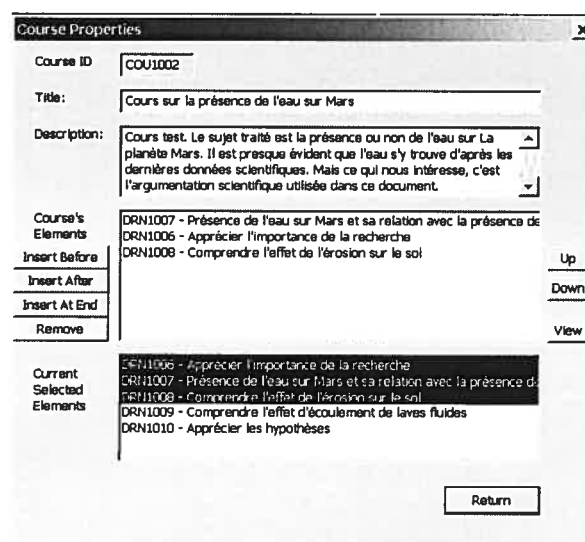


Figure 6-14 – Éditeur de définition de cours

Le responsable de la formation dispose de deux modes de sélection des éléments du modèle didactique : *le mode d'association large* et *le mode d'association ciblée*.

Dans *le mode d'association large* le responsable de la formation peut décider d'associer un ou plusieurs sujets génératifs pour répondre à un besoin donné. Une telle association se traduit sur le plan pédagogique par la planification dans le futur cours des éléments suivants:

- l'étude du sujet génératif associé,
- l'étude des sous-sujets qui lui sont associés,
- la réalisation des objectifs de compréhension qui leur sont rattachés, et
- la lecture des documents ou des parties de documents qui permettent la réalisation de ces objectifs de compréhension.

À cause de cette décomposition, le nombre et le volume du matériel à lire pour l'étude d'un sujet génératif donné peuvent être relativement élevés. C'est pourquoi on parle de *mode d'association large* (Figure 6-15). Ce mode est utile quand le responsable de la formation veut couvrir le maximum des sujets liés à un besoin de formation pour s'assurer que rien ne manque pour répondre à ce besoin. Seulement cette façon de faire comporte le risque de rendre le cours inutilement lent si des éléments non nécessaires sont planifiés. Et ce risque augmente de plus en plus à mesure que les sujets génératifs choisis sont placés vers le haut de la hiérarchie des sujets génératifs dans le modèle didactique du système (Figure 6-15).

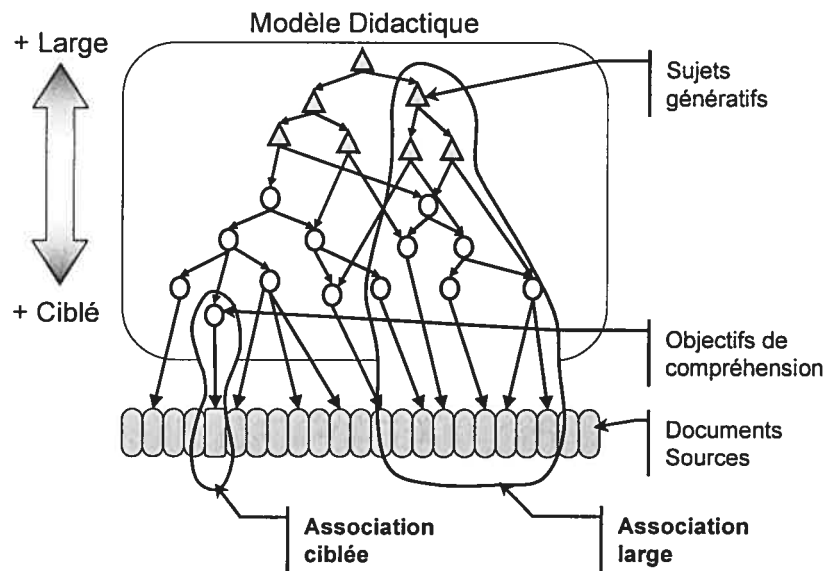


Figure 6-15 – Mode d'association large et ciblée

Le *mode d'association ciblée* est l'autre mode d'association offert par le système pour éviter le risque de mettre trop de choses dans un cours. Dans ce mode, le responsable de la formation choisit directement les objectifs de compréhension à atteindre au lieu de choisir les sujets génératifs. Le matériel pédagogique à lire pour réaliser un objectif de compréhension ciblé est relativement peu volumineux par rapport à ce qui est nécessaire pour un sujet génératif. Cependant, à cause de cette spécialisation, on a le risque que l'élément choisi ne réponde pas exactement ou complètement au besoin spécifié. Et ce risque augmente de plus en plus à mesure que l'objectif de compréhension choisi est placé vers le bas de la hiérarchie des objectifs de compréhension dans le modèle didactique du système (Figure 6-15).

L'ensemble des objectifs de formation $ObjFor()$ est défini comme suit :

$$ObjFor() = \langle ObjFor_Large(), ObjFor_Ciblé() \rangle \text{ ou}$$

$ObjFor_Large() \subset \mathcal{P}(sG)$ est l'ensemble des sujets génératifs retenus,

$ObjFor_Ciblé() \subset \mathcal{P}(oC)$ est l'ensemble des objectifs de compréhension retenus.

$\mathcal{P}(sG)$ et $\mathcal{P}(oC)$ sont les sujets génératifs et les objectifs de compréhension.

6.1.6 L'approche Document-based

Dans la pratique il arrive souvent que les documents didactiques à utiliser soient déjà fixés au préalable pour différentes raisons. Dans ce cas, l'approche *Document-based* permet au responsable de la formation de commencer la préparation d'un cours par la sélection du document didactique à utiliser.

L'opération de sélection du contenu dans l'approche *Document-based* peut se faire à différents niveaux de granularité par marquage direct sur le document source (petite sélection ou large sélection). En fonction des séquences sélectionnées, le système *IT-EKC* utilise les trois modèles de connaissance (discours, épistémique et didactique) relatifs aux documents utilisés, pour déterminer les objectifs de compréhension traités dans le contenu sélectionné ainsi que les sujets génératifs concernés. Le résultat obtenu par le système est présenté au responsable de la formation pour vérification et éventuellement pour modification comme dans l'approche *model-based*.

L'algorithme suivant permet de retrouver les objectifs de compréhension en fonction du texte sélectionné dans un document didactique.

Algorithme *ObjFor Document based()*:

En entrée :

$\mathcal{T}[c_1 c_2 \dots c_n]$ une représentation du document didactique, $c_{i=1..n}$ est un caractère.
 $t_Selection$ l'ensemble des séquences de texte sélectionnées (sans chevauchement)

En sortie :

$oC_Selection$ la liste des objectifs de compréhension retenus

- Soit $t_Selection = \{ s_1, s_2, \dots, s_m \mid s_{i=1..m} \subset T \text{ et } \forall i \neq j, s_i \cap s_j = \emptyset \}$
- Déterminer $m\mathcal{A}_Selection$ l'ensemble des micropropositions atomiques du modèle du discours qui chevauchent $t_Selection$:

$$m\mathcal{A}_Selection = \{ m\mathcal{A} \in T \mid \exists s_i \in t_Selection \text{ et } m\mathcal{A} \cap s_i \neq \emptyset \}$$
- Soit $to_process = \emptyset$ une liste d'éléments épistémiques de types variés

- Initialiser $to_process \leftarrow mA_Selection$ et
 $oC_Selection \leftarrow \emptyset$
- Tant que $to_process \neq \emptyset$ faire :
 - Considérer e le premier élément de $to_process$
 - Retirer e de $to_process$
 - Si e est de type microproposition atomique (mA) ou micropropositions composées (mC) alors :
 Ajouter à $to_process$ toutes les micropropositions composées, les macropropositions et les formes épistémiques qui réfèrent à e .
 - Si e est de type macroproposition (Ma) alors :
 Ajouter à $to_process$ toutes les autres macropropositions et les formes épistémiques qui réfèrent à e .
 - Si e est de type forme épistémique (FE) alors :
 Ajouter à $to_process$ toutes les autres formes épistémiques qui réfèrent à e .
 Ajouter à $oC_Selection$ tous les objectifs de compréhension qui ont une activité épistémique basée sur la forme épistémique e
- Retourner $oC_Selection$

6.1.7 La spécification du public cible

La caractéristique principale qui nous intéresse dans le public qui va suivre une formation est le niveau de sa connaissance initial, $NivPC(public)$, défini par l'ensemble des objectifs de compréhension qu'il maîtrise déjà. Deux situations se présentent : soit le système dispose déjà de cette connaissance dans les modèles des apprenants qui composent ce public, soit cette connaissance n'est pas disponible, et dans ce cas on doit faire une estimation pour le niveau initial (ou le niveau d'entrée) souhaité chez le futur public qui veut suivre ce cours. Le responsable du cours utilise l'éditeur du public cible pour définir ces paramètres (Figure 6-16).

$$\text{NivPC}(\text{public}) = \{ oc \in \mathcal{P}(oc) \mid \text{public maîtrise ou supposé maîtrisé } oc \}$$

$\mathcal{P}(oc)$ est l'ensemble des objectifs de compréhension du modèle didactique

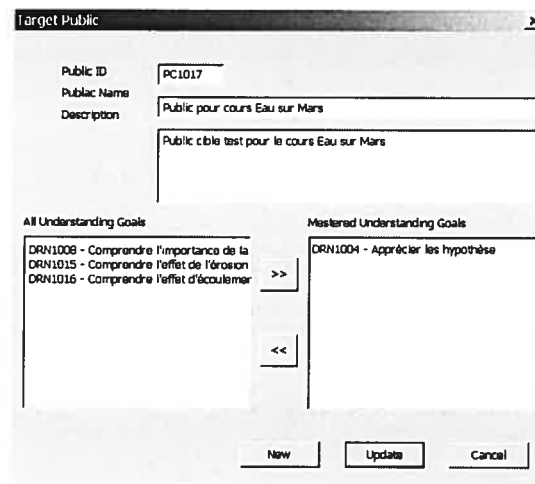


Figure 6-16 – Éditeur de public cible

6.1.8 Personnalisation des objectifs de formation

Les objectifs de formation définis initialement dans $ObjFor()$ sont définis d'une façon générale et ne tiennent pas compte du niveau de connaissance réel spécifique à chaque public cible. Un objectif de compréhension par exemple, peut être nécessaire à planifier pour un public donné et ne pas l'être pour un autre. Grâce à la définition du public cible $\text{NivPC}(\text{public})$, le module planificateur du système *IT-EKC* est capable de calculer les objectifs réels de formation adaptés à chaque public cible (ou à chaque apprenant, quand le groupe n'est constitué que d'un seul apprenant).

L'algorithme suivant permet de calculer les objectifs de formation personnalisés, $ObjForPerso()$, sous la forme d'un ensemble minimal d'objectifs de compréhension à réaliser, suffisant pour atteindre les besoins de formation initiaux. De plus cet algorithme convertit les objectifs de formation exprimés en sujets génératifs contenus dans $ObjFor_Large()$ en leurs équivalents exprimés en objectifs de compréhension.

L'algorithme reçoit en entrée les objectifs de formation initiaux $ObjFor()$ et le niveau du public cible $\text{NivPC}(\text{public})$ et retourne une liste d'objectifs de compréhension $listeOcPerso$, adaptés à ce public.

Algorithme *ObjForPerso()*: Calculer des objectifs de formation personnalisés

En entrée :

$ObjFor() = \langle ObjFor_Large(), ObjFor_Ciblé() \rangle$

$NivPC(public)$

En sortie :

$ListeOcPerso$

- Soit la liste des sujets génératifs $ListeSg = ObjFor_Large()$, initialisée à la liste des sujets génératifs initialement retenus
- Prendre en compte la dépendance entre les sujets : Pour chaque sujet génératif de $ListeSg$, insérer dans $ListeSg$ ses sujets génératifs pré-requis
- Prendre en compte la composition des sujets : Décomposer chaque sujet dans $ListeSg$ en ses sous-sujets et remplacer le par cette décomposition. Après cette étape la liste $ListeSg$ ne contient que des sujets génératifs non décomposables
- Soit la liste des objectifs de compréhension $ListeOcPerso = ObjFor_Ciblé()$, initialisée à la liste des objectifs de compréhension initialement retenus
- Pour chaque sujet s de la liste des sujets génératifs $ListeSg$, ajouter dans $ListeOcPerso$ l'ensemble des objectifs d'apprentissage liés à s . Chaque élément n'est inséré qu'une seule fois (pas de double)
- Épurer la liste des objectifs $ListeOcPerso$: les objectifs de compréhension déjà réalisés par le public cible ne doivent pas être replanifiés. Donc éliminer de $ListeOcPerso$ tous les objectifs présents dans $ListeOcPerso$ et dans $NivPC(public)$
- Prendre en compte la dépendance entre les objectifs de compréhension : Pour chaque objectif de $ListeOcPerso$, insérer ses objectifs de compréhension pré-requis
- Prendre en compte la composition des objectifs de compréhension : Décomposer chaque objectif dans $ListeOcPerso$ en ses sous objectifs et remplacer

le par cette décomposition. Après cette étape la liste *ListeOcPerso* ne contient que des objectifs de compréhension non décomposables

- Épurer la liste des objectifs *ListeOcPerso* de nouveau, car les deux étapes précédentes peuvent insérer des objectifs déjà atteints par le public cible
- Ordonner la liste des objectifs *ListeOcPerso* : si on considère les relations de pré-requis entre les objectifs de compréhension, ces objectifs sont topologiquement ordonnés. La liste $ListeOcPerso = [o_1, o_2, \dots, o_n]$, doit être ordonnée telle que:

$$\forall o_i \in [1..n] \text{ et } \forall o_j \in [1..n]:$$

Si o_i est un pré-requis de o_j , Alors $i < j$

6.1.9 Détermination du contenu à lire en fonction des objectifs de compréhension

À l'issue de la phase de personnalisation des objectifs de formation, nous disposons de la liste d'objectifs de compréhension à réaliser dans le cours (*ListeOcPerso*) et qui sont adaptés spécifiquement au public cible (*public*). Mais comme la liste *ListeOcPerso* ne contient que des objectifs de compréhension, elle ne spécifie que le “*quoi*” de la formation, c'est à dire sur *quoi* va porter la formation, sans spécifier le “*comment*” : quel contenu doit lire l'apprenant dans les sessions de lecture et quelles activités épistémiques doit-il réaliser lors de cette lecture.

Le but de cette étape, qui est la responsabilité du module planificateur, est de déterminer les séquences de texte à lire pour répondre aux objectifs de compréhension fixés. La procédure simple pour trouver ces séquences est la suivante :

Algorithme *SecLecture()*: Détermine l'ensemble des séquences de lecture S

Pour chaque objectif de compréhension O retenu dans la liste *ListeOcPerso* faire :

Retrouver l'activité épistémique AE liée à O

Retrouver la forme épistémique FE liée à AE

Décomposer récursivement la forme FE en ses sous formes jusqu'à arriver à des micropropositions atomiques $m\mathcal{A}$

Pour chaque microproposition atomique $m\mathcal{A}$ trouvée faire :

Retenir pour lecture la séquence de texte s référé par $m\mathcal{A}$

Inclure s dans S

La projection des éléments de S sur le document didactique (par changement de la couleur d'arrière plan), montre à l'apprenant ce qu'il doit lire précisément pour réaliser les d'objectifs de compréhension fixés (Figure 6-17). Comme les séquences de texte dans un document n'ont pas toutes la même importance, nous avons introduit deux modes de projection: un *mode de projection simple* et un *mode de projection multi-niveaux*. Dans le premier mode, le texte à lire est marqué par une seule couleur de fond et il n'y a pas de distinction entre les différentes parties du texte. Dans le mode de projection multi-niveaux la couleur du fond prend une teinte foncée proportionnellement à l'*importance* du texte. L'importance d'une séquence s , $Importance(s)$, est calculée simplement par la somme des poids heuristiques des formes épistémique qui réfèrent directement ou indirectement à cette séquence s .

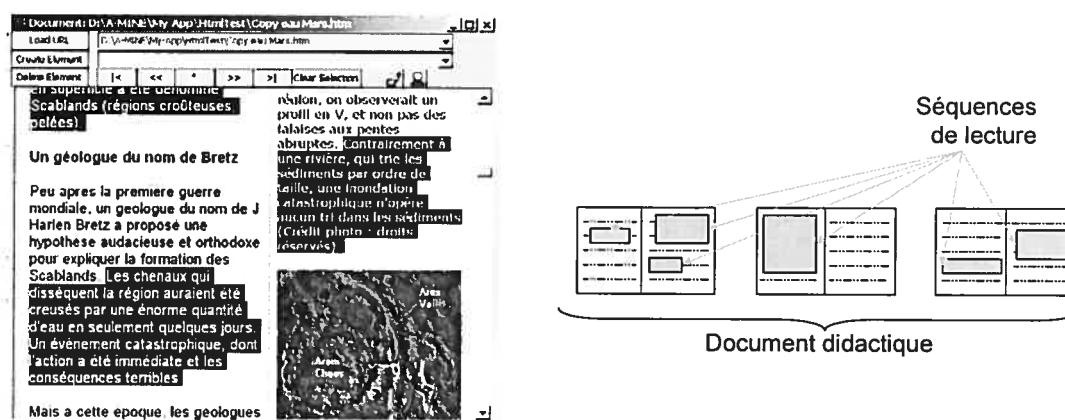


Figure 6-17 – Projection de S sur le document didactique

À des fins de simplification dans l'algorithme *SecLecture()* on a supposé qu'il n'y a qu'une seule activité épistémique pour atteindre un objectif de compréhension donné, et donc une

seule forme épistémique et un seul document à utiliser. Or, il arrive souvent que plusieurs documents différents permettent d'atteindre le même objectif de compréhension par des formes et des activités épistémiques différentes (O2, Figure 6-18). D'où se pose le problème de comment choisir entre ces alternatives. Par exemple dans l'illustration de la Figure 6-18, il est préférable de choisir l'activité A2.2 et son document Document_1 pour atteindre l'objectif de compréhension O2, car le même document est impliqué dans l'étude de l'objectif O1. Mais la situation devient rapidement compliquée quand le nombre de choix et d'alternatives devient important. Ce qui nécessite le recours à certaines stratégies pour effectuer le choix entre ces alternatives.

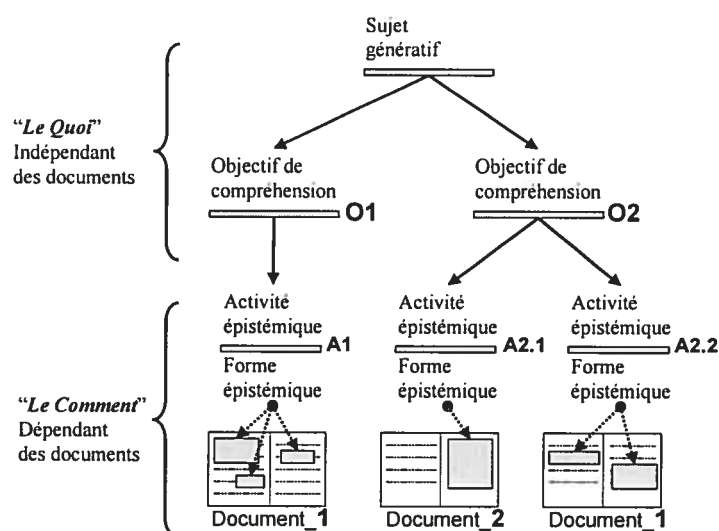


Figure 6-18 – Le “Quoi” et le “Comment” dans un cours

Nous avons identifié quatre stratégies pour faire la sélection entre les documents :

- **Favoriser les documents à référencement élevé** : dans cette stratégie on donne la préférence aux documents qui traitent le plus d'objectifs de compréhension possibles (référencement élevé). Cette méthode nécessite deux parcours de la liste des d'objectifs de compréhension *ListeOcPerso*. Dans le premier parcours on calcule le *degré de référencement* pour chaque document, $Dr(\text{document})$, c'est à dire le nombre d'objectifs de compréhension traités dans ce document. Dans le deuxième parcours on fait la sélection des documents : chaque fois qu'il y a des alternatives, nous choisissons le document qui a le *degré de référencement* le plus élevé.

L'avantage de cette stratégie est qu'elle minimise le nombre de documents différents à utiliser dans le cours (le même document y revient plusieurs fois). Par contre elle ne tient pas en compte des facteurs comme le volume de lecture ou la familiarité avec les documents.

- **Favoriser les documents minimisant le volume de lecture** : dans cette stratégie nous cherchons à minimiser le volume de lecture, pour des contraintes sur la durée de la formation par exemple. Cette méthode ne nécessite qu'un seul parcours de la liste des d'objectifs de compréhension *ListeOcPerso*. Pour chaque objectif de compréhension O qui peut être atteint par plus d'un document, nous calculons par document le *volume de lecture*, $Dv(\text{document}, O)$, qui est une métrique basée sur le nombre de caractères¹ des séquences de texte à lire dans *document* pour réaliser O . Le document à retenir est le document qui minimise la métrique $Dv(\text{document}, O)$.

$Dv(\text{document}, O)$:

Retrouver l'activité épistémique \mathcal{AE} qui est liée à O et qui utilise le document *document*

Décomposer récursivement la forme \mathcal{FE} en ses sous formes jusqu'à arriver à des micropropositions atomiques $m\mathcal{A}$ de *document*

$S \leftarrow 0$

Pour chaque microproposition $m\mathcal{A}$ trouvée faire :

$S \leftarrow 0 + (\text{la taille de la séquence de texte référé par } m\mathcal{A})$

Retourner S

L'avantage de cette stratégie est qu'elle minimise le volume de la lecture à faire par l'apprenant. Son inconvénient est le morcellement éventuel de la lecture sur plusieurs documents (à l'inverse de la stratégie précédente favorisant les documents à référencement élevé).

- **Favoriser les documents familiers** : à l'opposé des deux stratégies que nous venons de présenter qui utilisent des caractéristiques intrinsèques aux documents, cette stratégie utilise plutôt une caractéristique extrinsèque qui est la familiarité de l'apprenant avec le

¹ Le nombre de mots peut aussi servir dans cette métrique

document. On a défini la familiarité $\mathcal{D}f(\text{document})$, comme étant une fonction qui associe à chaque document un entier dans l'intervalle $[0,10]$: 0 pour un document totalement nouveau pour l'apprenant et 10 pour un document qui lui est très familier. Deux façons sont possibles pour déterminer cette valeur : c'est l'apprenant qui signifie lui-même son degré de familiarité avec les documents disponibles (c'est *la familiarité arbitraire*) ou c'est le système qui la calcule automatiquement en fonction de l'historique des utilisations des documents (c'est *une familiarité calculée*).

Remarquons que la notion de familiarité est très étroitement liée à la notion de "préférence", car plus un document est utilisé, et plus on peut présumer qu'il est préféré.

- **Choix aléatoire** : les trois stratégies exposées ci-haut ont un inconvénient qui peut devenir un vrai problème. C'est qu'elles favorisent l'utilisation d'un groupe de documents au détriment des autres. Or il peut y avoir dans les documents non utilisés d'excellents documents (parfois nouveaux) qui méritent la considération. Pour cette raison nous avons introduit une quatrième stratégie basée sur le choix *aléatoire* des documents où la sélection ne tient compte d'aucun critère que de celui de répondre aux objectifs de formation. Le responsable de la formation peut de temps à l'autre activer cette stratégie pour créer un genre de brassage entre les documents et donner ainsi une chance égale à tous les documents.

6.1.10 Stratégies d'exploration du contenu

Après l'étape de détermination du contenu qui a désigné les documents et les séquences de texte à lire pour atteindre les objectifs fixés, l'apprenant est en mesure de commencer sa lecture.

Pour s'accommoder aux différents styles de lecture des apprenants, dont certains aiment être très guidés dans leur apprentissage alors que d'autres aiment avoir plus de liberté, le système *IT-EKC* offre aux apprenants deux stratégies d'utilisation : la *stratégie activity-based* et la *stratégie discours-based* (Figure 6-19).

Dans la *stratégie activity-based*, l'activité pédagogique et l'ordre des lectures sont complètement contrôlés par le système. L'ordre des lectures est dicté par l'ordonnement des activités déterminé par le système. À chaque étape, le système sélectionne l'objectif à réaliser et

détermine la séquence de texte à lire pour y arriver. Ce qui fait que la lecture peut ne pas être dans l'ordre habituel. Des sauts en avant et en arrière dans le même document ou entre documents peuvent éventuellement survenir. L'avantage de cette stratégie est qu'elle garantit que les pré-requis essentiels pour la compréhension d'un passage sont planifiés et réalisés avant la lecture de ce passage.

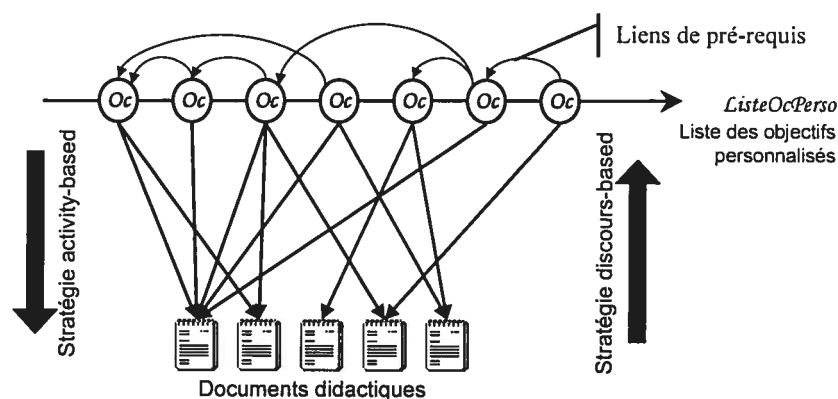


Figure 6-19 – Stratégie de présentation de contenu

À l'inverse, dans la *stratégie discours-based* l'ordre de lecture est laissé au choix de l'apprenant qui peut lire les documents dans l'ordre qu'il préfère (généralement dans l'ordre du discours habituel). Le système suit l'apprenant dans sa lecture et vérifie à chaque fois si le contenu affiché est relié à un objectif de compréhension à réaliser. Si c'est le cas, il enclenche l'activité épistémique associée et demande à l'apprenant de la réaliser. Le système contrôle aussi que les pré-requis nécessaires sont déjà réalisés, dans le cas contraire il conseillera à l'apprenant de lire le contenu relatif à ces pré-requis avant de continuer.

6.2 Stratégies d'interaction avec le système *IT-EKC*

L'apprenant peut utiliser le système *IT-EKC* de trois manières : par une utilisation directe, par une utilisation contrôlée ou par une utilisation libre. La désignation des modes possibles est la responsabilité du concepteur de cours.

6.2.1 Utilisation directe

Le mode l'utilisation directe permet d'exploiter les modèles de connaissance *LEKC* comme des ressources supplémentaires aux documents didactiques. L'apprenant peut les consulter

en ligne ou les imprimer. Le rôle du système est de fournir à l'apprenant une aide épistémique contextuelle : les interactions avec les éléments actifs¹ sur chaque passage du texte font apparaître les formes épistémiques qui lui sont reliées. La réflexion sur le sens de ces formes donne à l'apprenant l'occasion de canaliser et d'approfondir sa compréhension.

L'apprenant dispose d'un *bloc note interactif* pour la prise des annotations (Figure 6-20). Cette opération est grandement facilitée par le système. À la lecture d'un passage de texte où l'apprenant sent le besoin de prendre une note, l'apprenant peut créer une entrée dans son bloc note interactif et ce de deux manières:

- inscrire son propre texte dans l'entrée comme dans la prise de notes conventionnelles et qui sera associé au passage de texte courant, ou plus simple encore,
- copier et coller dans le bloc note un élément, une partie ou la totalité d'une forme épistémique reliée à ce passage.

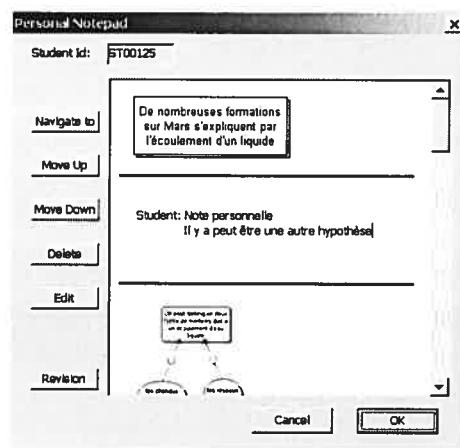


Figure 6-20 – Bloc note interactif

La suite des entrées dans le bloc note interactif, donne à l'apprenant une vision de synthèse du document didactique et peut servir comme outil de révision. De plus, grâce à la fonction de navigation (représentée par le bouton “*Navigation*”), l'apprenant peut retracer l'origine de chaque entrée dans cette suite, soit directement dans le document source pour les micropropositions ou soit indirectement par la hiérarchie des formes épistémiques pour les autres éléments.

¹ les méta-tags

6.2.2 Utilisation contrôlée

L'utilisation contrôlée comme nous l'avons définie au chapitre 5 est une utilisation qui entre dans le cadre d'un cours à suivre. Le module tuteur (Figure 6-21) parcourt la suite des activités épistémiques planifiées pour ce cours et demande à l'apprenant de les exécuter, c'est à dire de construire ou de compléter les formes épistémiques correspondantes à la lecture du texte.

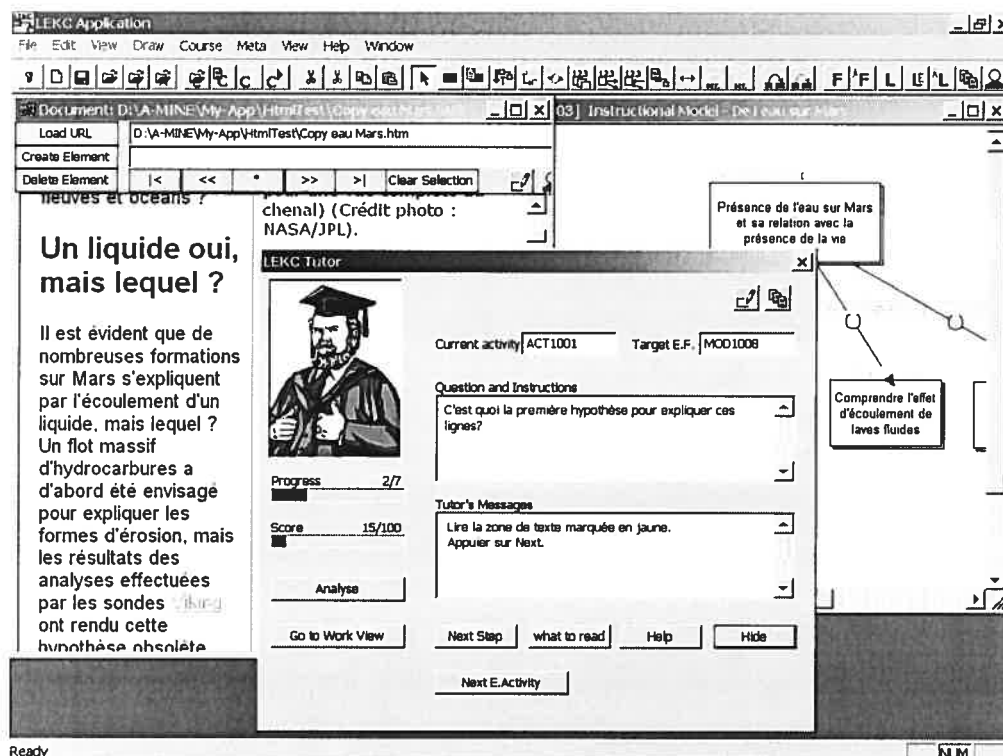


Figure 6-21 – Interface tuteur *IT-EKC*

L'ordre de construction

Pour la construction de chaque élément (nœud ou lien) d'une forme épistémique donnée, le module tuteur marque le texte correspondant et présente à l'apprenant les instructions et les indices nécessaires à sa tâche. L'ordre des éléments dans la construction a été préalablement spécifié par le concepteur et tente de faciliter le travail de l'apprenant. Une stratégie que nous suggérons pour cet ordre est d'ordonner les éléments à construire en fonction de leurs positions dans la pyramide de la connaissance du domaine, introduite au niveau du chapitre trois (Figure 3-8) : les premiers éléments à construire sont ceux à la base de cette pyramide

(comme les connaissances structurelles ou les faits) et les derniers éléments à construire sont au sommet (comme les principes et les idées complexes). Une contrainte générale à respecter est que les éléments épistémiques décrivant des relations (type lien) sont à construire après la construction des éléments qu'ils mettent en relation.

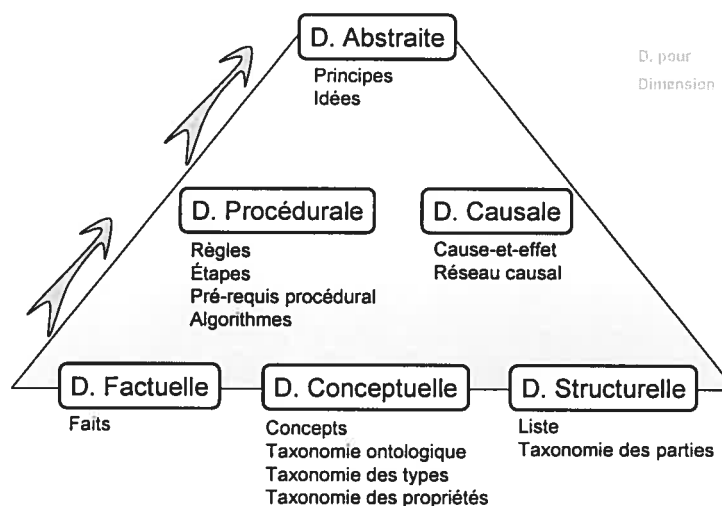


Figure 6-22 – Pyramide de la connaissance du domaine

Responsabilités du module tuteur et stratégies d'intervention

Le module tuteur suit à la trace le travail de l'apprenant et observe ses interactions avec l'éditeur des formes épistémiques. En fonction de ces interactions, le système détermine quand et comment intervenir pour assister l'apprenant. Nous avons mis au point trois stratégies d'intervention pour le système : *Stratégie Model-tracing simple*, *Stratégie Model-tracing avancé* et *Stratégie Configuration-based*

Mais avant de discuter de ces stratégies, notons que pragmatiquement l'observation du système pour les interactions de l'apprenant se limite à deux types d'interactions élémentaires:

- **Opération d'ajout** : L'apprenant ajoute un élément épistémique à la forme en construction. L'élément peut être un nœud ou un lien d'un type donné. L'ajout simultané d'un ensemble de nœuds et de liens (dans les opérations copier-coller) peut être considéré comme l'ajout individuel de chaque élément de cet ensemble.

- **Opération de retrait** : L'apprenant retire un élément épistémique de la forme en construction.

Le travail de l'apprenant dans la construction d'une forme épistémique se résume donc à une suite d'opérations d'ajout et de retrait d'éléments épistémiques. La stratégie d'intervention du système est basée sur cette succession d'opérations.

6.2.2.1 Stratégie du Model-tracing simple

La stratégie du model-tracing simple est une stratégie d'intervention du système qui implique une réaction immédiate vis-à-vis des interactions de l'apprenant avec l'interface. Le système dispose de la forme épistémique correcte à construire et réagit immédiatement quand l'apprenant dévie par rapport à cette forme, soit par ajout d'éléments étrangers soit par retrait ou oubli d'éléments corrects. La stratégie est dite simple car le concepteur n'a à définir que la forme épistémique jugée correcte sans avoir à définir les formes épistémiques incorrectes (*bugs library*).

Soit $\mathcal{F}\mathcal{E}$ la forme épistémique à construire par l'apprenant à une étape donnée¹. $\mathcal{F}\mathcal{E}$ est un graphe $\langle \mathcal{N}, \mathcal{A} \rangle$ où \mathcal{N} est l'ensemble des nœuds et où \mathcal{A} est l'ensemble des liens. Soit $f\acute{e}\mathcal{A}$ la forme épistémique en construction par l'apprenant. Le but du système est d'amener l'apprenant à construire la même forme épistémique que celle construite par l'expert de la matière (le *SME*). C'est à dire que $f\acute{e}\mathcal{A} = \mathcal{F}\mathcal{E}$.

L'apprenant par le biais de l'éditeur des formes épistémiques ajoute et retire des éléments de $f\acute{e}\mathcal{A}$. En tenant compte de ces deux types d'interactions et du fait que les éléments utilisés peuvent être corrects ou non, le système considère quatre cas possibles : ajout d'un élément correct, ajout d'un élément étranger, retrait d'un élément correct et retrait d'un élément étranger.

1. Ajout d'un élément correct α : l'apprenant insère dans $f\acute{e}\mathcal{A}$ un élément épistémique correct α , $\alpha \in \mathcal{F}\mathcal{E}$. À ce stade l'apprenant ne sait pas si cet élément est nécessairement correct ou pas. Le système a deux stratégies de réaction possibles :

Stratégie de renforcement : si l'apprenant semble éprouver des difficultés dans son

¹ $\mathcal{F}\mathcal{E}$ a été initialement construite par l'expert de la matière

avancement ou manquer de confiance¹, le système donne alors des messages d'encouragement et de félicitation pour augmenter la confiance de l'apprenant. Ces messages (du genre "C'est excellent...", "Bien, ...") sont pondérés en fonction du poids épistémique de l'élément ajouté.

Stratégie de perturbation : si l'apprenant semble être sûr de lui et avance avec un bon rythme, le système peut ne pas intervenir (avec des messages d'encouragement comme dans le cas du renforcement) ou peut demander à l'apprenant une explication pour son action. Les deux situations mettent l'apprenant dans le doute et le poussent à repenser son action, ce qui le force à réfléchir davantage.

2. Ajout d'un élément étranger α : l'apprenant insère dans $f\mathcal{E}\mathcal{A}$ un élément épistémique étranger α , $\alpha \notin \mathcal{F}\mathcal{E}$. Nous disons "étranger" au lieu de "erroné" car le système n'a pas la capacité, ni la connaissance nécessaire pour juger qu'un élément qui n'est pas dans $\mathcal{F}\mathcal{E}$ est nécessairement erroné. Pour cette raison il demande à l'apprenant de donner une explication à son action. Le système vérifie si le même élément étranger a été inséré par d'autres apprenants dans le même contexte², auquel cas il affiche à l'apprenant la liste d'explications déjà données (triée par ordre décroissant sur le nombre d'utilisations) et l'apprenant en sélectionne une qui le satisfait ou en ajoute une nouvelle qui est mémorisée à son tour.
3. Retrait d'un élément correct α : l'apprenant retire de $f\mathcal{E}\mathcal{A}$ un élément épistémique correct α , $\alpha \in \mathcal{F}\mathcal{E}$. Ceci montre que l'apprenant n'était pas sûr de son action d'ajout de cet élément. Pour attirer l'attention de l'apprenant, le système indique à l'apprenant (par marquage) sur le document didactique, la ou les séquences de texte qui suggère que α soit un élément correct et lui demande de la ou les relire plus attentivement. Si après cette lecture l'apprenant insiste sur son action de retirer l'élément α , le système lui demande alors de donner ou de sélectionner une explication pour la justifier.

¹ par analyse de l'historique de ses actions d'ajout et de retrait

² dans la même activité épistémique

4. Retrait d'un élément étranger α : l'apprenant retire de $f\mathcal{EA}$ un élément épistémique étranger α , $\alpha \notin \mathcal{FE}$. L'apprenant s'est probablement rendu compte que cet élément n'est pas correct dans ce contexte. Le système demande à l'apprenant une explication sur ce geste. Cette explication sur le retrait de l'élément α est en principe en contradiction avec l'explication donnée pour son insertion.

L'utilisation de la stratégie du model-tracing simple dans la construction d'une forme épistémique et sur plusieurs cycles de cours avec un grand nombre d'apprenants permet au système de ramasser beaucoup d'informations sur :

- Le degré de difficulté pour chaque élément épistémique : cette métrique est calculée par le nombre d'opérations infructueuses d'ajouts et de retraits d'éléments avant que le bon élément ne soit bien placé. Cette métrique peut être améliorée par la prise en compte du facteur temps (la durée que prend en moyenne cette opération).
- La nature des erreurs commises (d'ajouts d'éléments étrangers et de retraits d'éléments corrects)
- Les explications données par les apprenants pour leurs actions, tant pour les erreurs que pour leurs corrections.

L'intérêt de réunir toute cette information est double. Premièrement le concepteur de cours peut la consulter et l'analyser périodiquement dans le cadre de la maintenance continue du matériel didactique. Il peut en conséquence modifier le contenu ou réorganiser les formes épistémiques concernées pour éviter les difficultés relevées. Deuxièmement, les interventions du système peuvent être améliorées par la prise en compte de cette information dans les dialogues avec l'apprenant comme nous allons le voir dans la stratégie du model-tracing avancé.

6.2.2.2 Stratégie du Model-tracing avancé

La stratégie du model-tracing avancé a le même principe que la stratégie du model-tracing simple. Mais elle se différencie par deux caractéristiques : l'utilisation d'une librairie d'erreurs et la génération de dialogues pseudo-socratiques.

Utilisation d'une librairie d'erreurs

L'information réunie par l'utilisation de la stratégie du model-tracing simple permet au concepteur de cours d'identifier les erreurs types commises par les apprenants. Ces erreurs sont de deux natures : 1) des erreurs d'ajout d'éléments étrangers et 2) des erreurs de retrait d'éléments corrects. Pour chaque erreur définie le concepteur prévoit des explications et éventuellement des indices que le système utilisera dans son intervention à l'apparition de cette erreur.

Génération de dialogues pseudo-socratiques

Les dialogues socratiques ont été longtemps utilisés comme outils d'apprentissage. Leur principe est de dialoguer avec l'apprenant pour le pousser à corriger ses propres erreurs de raisonnement et ceci en cherchant à le mettre en contradiction. Implémenter les dialogues socratiques sur un système informatisé est extrêmement complexe car il exige des capacités très évoluées comme le raisonnement et le traitement du langage naturel qui sont très difficiles à mettre en œuvre dans un cadre pratique.

Pour éviter cette complexité inhérente aux dialogues socratiques, nous avons introduit dans notre système *IT-EKC* un mécanisme d'intervention qui *simule* (dans une certaine mesure) un dialogue socratique avec l'apprenant. C'est la génération des dialogues pseudo-socratiques.

L'idée des dialogues pseudo-socratiques est basée sur l'exploitation de la base des explications qui s'est formée au cours de l'utilisation répétée du système. Ces explications sont fournies par les apprenants eux-mêmes et éventuellement par l'expert de la matière. Elles constituent une connaissance collective sous forme textuelle qui bien qu'elle soit indéchiffrable pour le système, lui permet de réagir aux interactions de l'apprenant¹ grâce aux relations qui existent entre ses éléments :

- La relation de *support* : une explication qui supporte l'insertion d'un élément épistémique,
- La relation d'*opposition* : une explication qui s'oppose à l'insertion d'un élément épistémique, et

¹ insertions et aux retraits d'éléments épistémiques

— La relation de *contradiction* : une explication qui contredit une autre explication.

Soit $\mathcal{PE_supporte}(\alpha)$ l'ensemble des explications qui supportent l'ajout de α dans \mathcal{PE} et soit $\mathcal{PE_oppose}(\alpha)$ l'ensemble des explications qui s'y opposent. Soit e une explication donnée qui supporte ou s'oppose à l'ajout de α . Nous définissons $\mathcal{PContr}(e)$ comme l'ensemble des explications qui contredisent l'explication e (Figure 6-23).

Si $e \in \mathcal{PE_supporte}(\alpha)$ alors $\mathcal{PContr}(e) \subset \mathcal{PE_oppose}(\alpha)$, et inversement

Si $e \in \mathcal{PE_oppose}(\alpha)$ alors $\mathcal{PContr}(e) \subset \mathcal{PE_supporte}(\alpha)$.

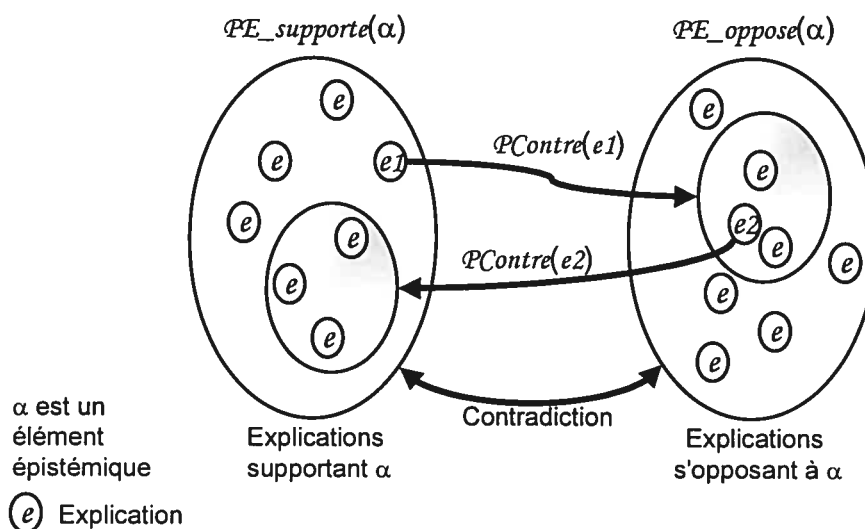


Figure 6-23 – Explication supportant / s'opposant à l'insertion de α

Quand la stratégie du dialogue pseudo-socratique est activée, le système réagit aux erreurs de construction commises par l'apprenant, par le scénario ci-après. Le cas traité est l'insertion d'un élément épistémique erroné α . Des éléments de détail ont été supprimés pour simplification.

Scénario de dialogue (insertion élément erroné α)

- Le système demande à l'apprenant de sélectionner une explication qui supporte l'insertion de α
- L'apprenant consulte la liste $\mathcal{PE_supporte}(\alpha)$ d'explications supportant l'insertion de α et sélectionne une explication $eA1 \in \mathcal{PE_supporte}(\alpha)$
- Tant que l'élément α est inséré et qu'il reste des explications non présentées faire :
 - Mettre l'apprenant en contradiction : le système présente à l'apprenant une explication $eS1$ s'opposant à l'insertion de α et contredisant l'explication $eA1$. $eS1 \in \mathcal{PE_oppose}(\alpha)$ et $eS1 \in \mathcal{PContre}(eA1)$.
 - Si l'apprenant accepte la nouvelle explication du système $eS1$ et reconnaît la contradiction dans son raisonnement, il réagit en retirant l'élément erroné α . Sinon, il indique au système une nouvelle explication $eA2$ qui supporte l'insertion de α et contredit l'explication $eS1$ du système. $eA2 \in \mathcal{PE_supporte}(\alpha)$ et $eA2 \in \mathcal{PContre}(eS1)$.
 - $eA1 \leftarrow eA2$ pour recommencer la boucle

La même logique du scénario est utilisée pour les cas d'erreurs de retrait d'éléments épistémiques corrects, en inversant le notion de *support* et d'*opposition* (Figure 6-24).

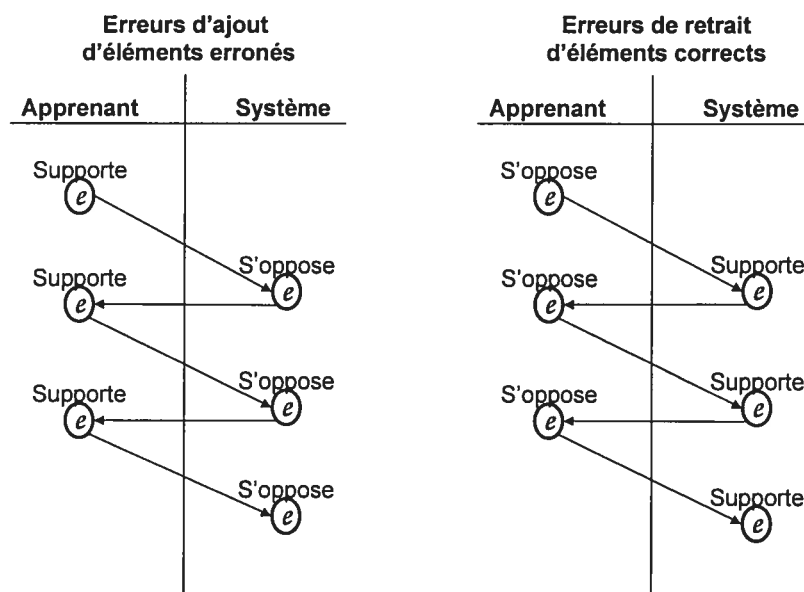


Figure 6-24 – Interaction apprenant / système dans un dialogue pseudo-socratique

6.2.2.3 Stratégie Configuration-based

Les deux stratégies présentées, de model-tracing simple et de model-tracing avancé, sont utilisées par le système pour réagir aux erreurs élémentaires, c'est-à-dire impliquant un seul élément épistémique (un nœud ou un lien, inséré ou retiré d'une forme en construction). Or il y a un autre type d'erreurs que ces deux stratégies ne permettent pas de traiter, ce sont les erreurs de configuration.

Les erreurs de configuration sont des erreurs relatives à l'organisation et à la structuration des éléments de la forme en construction. Les mêmes éléments dans une forme peuvent exprimer un sens différent s'ils sont disposés différemment. La stratégie configuration-based permet au système *IT-EKC* de réagir à l'apparition d'une certaine configuration d'éléments dans la construction de l'apprenant.

L'analyse de l'historique accumulé sur les interactions avec les apprenants dans les deux stratégies model-tracing simple et de model-tracing avancé permet au concepteur du cours d'identifier les erreurs de configuration types que les apprenants commettent et d'associer à chaque configuration des explications et des indices.

Cette librairie d'erreurs de configuration est utilisée par le système pour contrôler le travail de construction épistémique de l'apprenant. Dès que le système détecte et identifie une

configuration erronée, il met en évidence¹ les éléments qui la composent et initie des réactions semblables à la stratégie du model-tracing avancé, mais cette fois sur les configurations. Des explications de support et d'opposition à cette configuration sont utilisées par le système pour la génération de dialogues pseudo-socratiques sur les configurations.

Réaction différée

Une configuration C est définie par un sous-ensemble de nœuds et de liens dans une forme épistémique donnée \mathcal{EF} . $C \subset \mathcal{EF}$. Soit sC une partie de C . sC est aussi une configuration dans \mathcal{EF} . Nous disons que sC est une sous-configuration de C . $sC \subset C$.

Considérons le cas où C et sC sont erronées toutes les deux. Pour que C soit construite par l'apprenant, il faut qu'il construise sC en premier car sC est une partie de C . Ce qui fait que le système va toujours détecter sC avant de détecter C . Si le système réagit à l'apparition de sC et arrive à convaincre l'apprenant de la corriger, cette correction va faire disparaître sC et du même coup va faire disparaître C . Dans ces conditions l'apprenant n'aura jamais l'occasion *d'être confronté* à la configuration erronée C .

Pour éviter cette impasse nous avons introduit un mécanisme de *réaction différée*. Quand cette option est activée par le responsable de la formation, le système ne réagit plus immédiatement à l'apparition d'une configuration erronée sC si elle est une sous-configuration d'une configuration erronée C . La réaction du système à sC sera différée jusqu'après la réaction à C . Si l'apprenant avance dans sa construction mais sans jamais arriver à la configuration C , le système peut réagir à sC après une certaine période (fixée par le concepteur).

En plus de donner à la configuration C une chance d'arriver, ce mécanisme donne aussi à l'apprenant une chance de corriger sC de lui-même sans l'aide du système.

¹ changement de la couleur et de l'épaisseur du contour

6.2.2.4 Application des trois stratégies

Les trois stratégies de réaction du système *IT-EKC* que nous venons d'exposer ne sont pas antagonistes mais complémentaires. Elles peuvent être actives toutes les trois en même temps selon le degré de la connaissance disponible au système (Figure 6-25). La stratégie du model-tracing simple sera utilisée fréquemment car elle n'exige que la présence d'une forme épistémique de référence faite par le concepteur de cours. La stratégie du model-tracing avancé sera utilisée pour les cas d'erreurs prédéfinies dans la librairie d'erreurs, avec possibilité d'engager l'apprenant dans des dialogues pseudo-socratiques. Pour les erreurs de configuration, la stratégie configuration-based est mise en œuvre avec la possibilité de réactions différées et de dialogues pseudo-socratiques sur les configurations (c'est à dire des parties de la forme épistémique en construction).

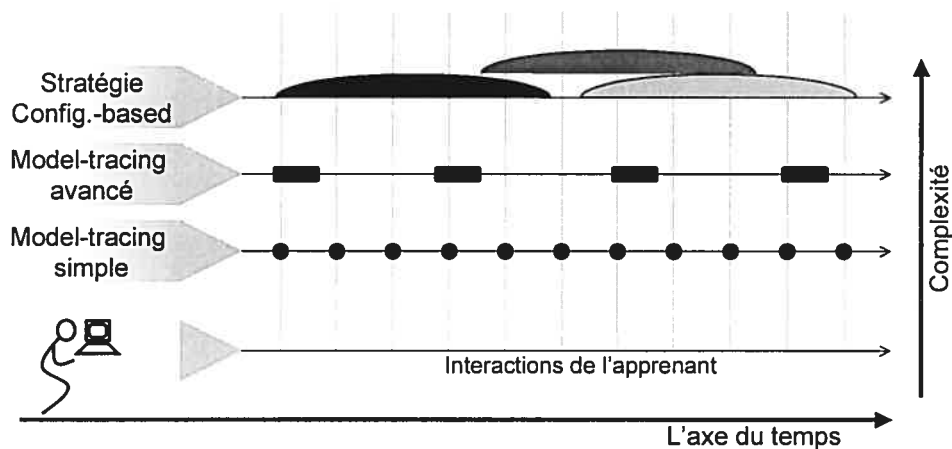


Figure 6-25 – Utilisation simultanée des trois stratégies d'intervention

Un cours n'est généralement pas opérationnel à 100% au premier cycle de son utilisation. Comme pour la programmation, un cours nécessite beaucoup de corrections et d'ajustements qui ne se montrent que dans l'utilisation du cours avec des apprenants dans un vrai contexte d'apprentissage.

Les différentes stratégies proposées ne peuvent pas toutes s'appliquer dès le premier jour de la création d'un cours (Figure 6-26). La stratégie du model-tracing simple peut être utilisée quand un cours est déployé pour la première fois. La stratégie du model-tracing avancé sera mise en œuvre après plusieurs cycles d'utilisation du cours et après l'analyse des erreurs types les plus fréquemment commises par les apprenants dans la construction des formes

épistémiques. La stratégie du configuration-based ne peut s'appliquer que plus tard quand assez d'informations sont réunies à propos des configurations d'éléments qui posent le plus de problèmes aux apprenants.

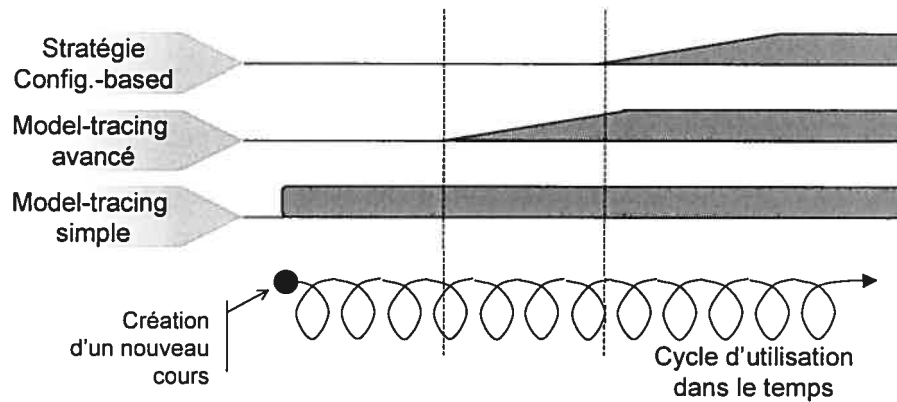


Figure 6-26 – Application des trois stratégies et cycles du cours

6.2.3 Utilisation libre

Le mode d'utilisation libre est prévu pour les bons apprenants qui arrivent à réaliser les différentes tâches épistémiques planifiées dans un cours et qui veulent exprimer leurs propres visions pour le sujet ou suggérer des améliorations à la conception de référence faite par le concepteur de cours.

Dans ce type d'utilisation les apprenants disposent dans l'environnement de travail des mêmes outils de conception que le concepteur de cours. Grâce aux nouvelles formes épistémiques que les apprenants créent, la connaissance du système s'étend et s'enrichit et les apprenants deviennent des *producteurs* de la connaissance et non simplement des *consommateurs*.

Le résultat de la modélisation libre de chaque apprenant constitue *le modèle cognitif personnel*, par opposition au *modèle cognitif contrôlé* où le système dicte les choses à réaliser en fonction du modèle épistémique de référence donné par le concepteur. L'addition et l'accumulation des travaux de modélisation épistémiques libres et contrôlés faits par les apprenants qui utilisent le système, en plus de la conception de référence faite par les experts de la matière, constituent une large *base de connaissance collective* qui est en évolution continue (Figure 6-27).

Le rôle du système est d'exploiter cette base de connaissance pour assister et superviser les apprenants dans leurs travaux de modélisation.

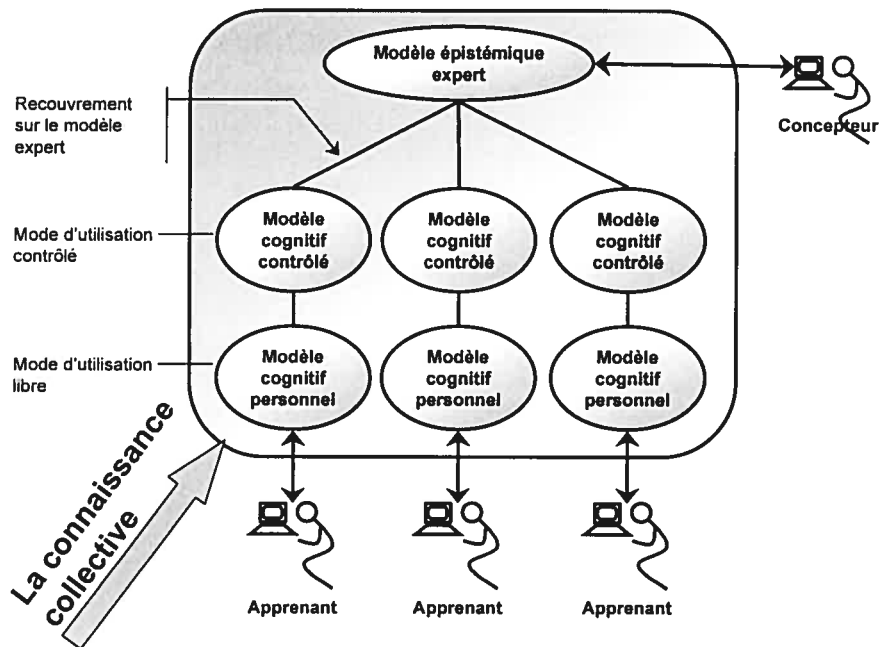


Figure 6-27 – Base de connaissance collective

La supervision des travaux de modélisation

Les nouvelles formes épistémiques créées par l'apprenant sont libres de toutes contraintes. D'une façon pratique, le système ne peut les contrôler que par comparaison aux autres formes épistémiques déjà mémorisées dans sa base de connaissance.

Le système tente continuellement d'identifier des configurations d'éléments dans le travail de l'apprenant qui s'apparentent à des configurations déjà mémorisées dans la base de connaissance collective (Figure 6-28). Quand l'appariement réussit et que la configuration reconnue présente des erreurs, le système enclenche la stratégie de *configuration-based* sur cette configuration pour conseiller l'apprenant. Il peut lui donner des indices, des explications ou initier avec lui un dialogue pseudo-socratique. Ce qui pousse l'apprenant à reprendre sa conception et à approfondir sa réflexion. Les contre-explications qu'il donnerait éventuellement, vont venir enrichir la capacité de dialogue du système.

Le processus d'appariement continu qu'effectue le système a une autre utilité. Quand le système reconnaît qu'un certain nombre de sous-configurations d'une configuration donnée C , ont été construites par l'apprenant, le système peut anticiper sur l'intention de l'apprenant de construire C et lui suggérer de compléter le reste de la configuration automatiquement. C'est un mécanisme d'assistance semblable à l'*auto-completion* utilisée dans les environnements de développement de logiciels que nous appliquons ici aux formes épistémiques (Figure 6-28). Ce mécanisme accélère considérablement le travail de l'apprenant quand les configurations qu'il souhaite construire ont été réalisées dans le passé par d'autres apprenants ou éventuellement par le concepteur.

Plusieurs configurations peuvent avoir en commun les mêmes sous-configurations, ce qui pose un problème d'ambiguïté pour décider quelle configuration l'apprenant cherche à construire. Dans cette situation, le système présente à l'apprenant les configurations possibles à partir de ces sous-configurations et lui suggère d'opérer une sélection.

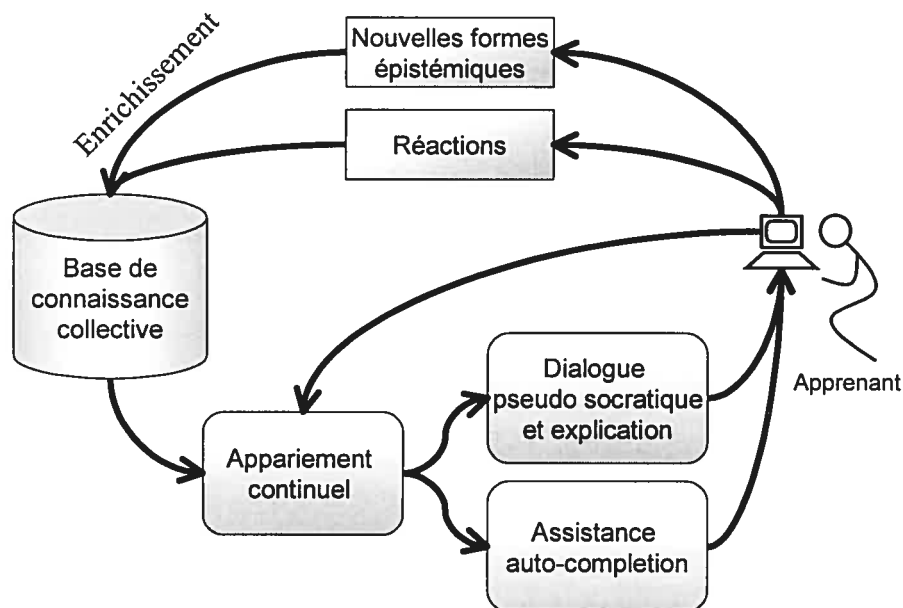


Figure 6-28 – Rôle du système dans la supervision

6.3 Conclusion du chapitre

Nous avons présenté dans ce chapitre le détail de la création et du déploiement d'un cours dans un système STI basé sur l'approche de modélisation *LEKC* et sur l'architecture *IT-EKC*.

Nous avons exposé à l'aide du prototype développé, la démarche à suivre par le concepteur de cours et les outils à utiliser dans le processus de modélisation, et nous avons détaillé les stratégies d'intervention du système dans les différentes situations qu'occasionnent les interactions avec l'apprenant.

CHAPITRE 7

DISCUSSION ET CONCLUSION

La plupart des STIs actuels utilisent la stratégie pédagogique de réparation d'erreur, dite *stratégie de remédiation* [Rouane, Frasson et Kaltenbach, 2002-a]. Cette stratégie est basée sur un cycle répétitif constitué de trois types d'activités pédagogiques (Figure 7-1):

- présentation du contenu non-évaluable : ce type d'activité est essentiellement lié à la présentation de la connaissance à apprendre, généralement sous forme textuelle ou de contenu multimédia (narration¹, vidéo, etc.) Dans ce contexte l'apprenant est récepteur de la connaissance (ou du message) véhiculée. À ce stade, aucun contrôle ou vérification de cette réception ne sont effectués.
- présentation du contenu évaluable : ce type d'activités est lié à la présentation des exercices de contrôle, qui ont pour rôle de vérifier le processus d'acquisition qui s'est déroulé dans l'étape précédente de présentation de contenu non-évaluable.
- prise de décisions pédagogiques : En fonction des résultats obtenus dans la présentation du contenu d'évaluation, les systèmes effectuent généralement des actions pour remédier aux problèmes et difficultés observés dans cette phase de contrôle.

Le déphasage dans cette stratégie de remédiation entre la phase de réception de la connaissance et la phase de contrôle de cette réception est à l'origine de beaucoup de

¹ Voix off

problèmes, dus essentiellement au fait que ces deux processus (la réception et le contrôle de la réception) se déroulent dans des contextes différents et différés dans le temps.

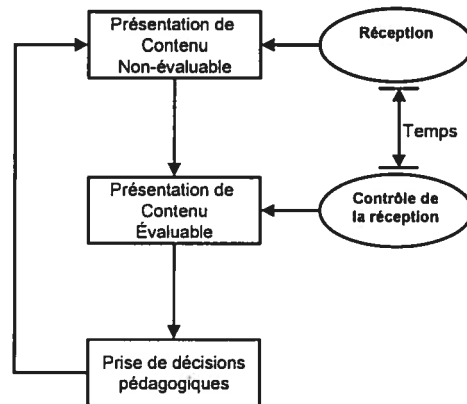


Figure 7-1 – Circuit pédagogique classique et la stratégie de remédiation

7.1 Inconvénient de la stratégie de remédiation

La stratégie de remédiation (ou de contrôle différé) a deux inconvénients majeurs : *la limite de la couverture des contrôles* et *la perte de contexte d'erreurs*.

La limite de la couverture des contrôles. La connaissance présentée dans la phase de lecture est généralement très chargée et comporte beaucoup de détail. Mais dans la phase de contrôle, toute cette connaissance ne peut pas faire l'objet de tests présentés à l'apprenant. Seulement une petite fraction (un échantillon) de cette connaissance, jugée importante, fera l'objet de contrôle. De ce fait, la plus grande partie de la connaissance présentée à l'apprenant dans la phase de la lecture reste incontrôlée, avec toutes les erreurs de compréhension que cela suppose.

La perte du contexte d'erreurs. Les erreurs de compréhension naissent et se développent généralement lors de la phase de lecture. Quand ces erreurs ne sont pas corrigées rapidement, elles donnent naissance à des erreurs de compréhension (secondaires), qui à leur tour donnent naissance à d'autres générations d'erreurs de compréhension. C'est le phénomène *d'erreurs en cascade* qui est une grande source de surcharge cognitive de l'apprenant [Rouane, Frasson et Kaltenbach, 2002-a]. Le contexte dans lequel une erreur de compréhension apparaît, en plus de l'enchaînement d'erreurs initiales qui l'ont provoquée, sont très importants à connaître pour pouvoir corriger cette erreur. Or avec la méthode du

contrôle différé de la stratégie de remédiation, les contextes d'erreurs ne sont ni suivis, ni mémorisés. Ce qui rend extrêmement difficile la correction des erreurs détectées.

7.2 Vers une stratégie de prévention

L'idée que nous proposons avec notre approche de lecture assistée *LEKC*, est de permettre que la phase de réception et la phase de contrôle de la connaissance, soient effectuées simultanément lors de la phase de lecture (Figure 7-2). Ceci nous fait passer d'une "stratégie de remédiation" pour les erreurs de compréhension, à une "stratégie de prévention" pour ces erreurs.

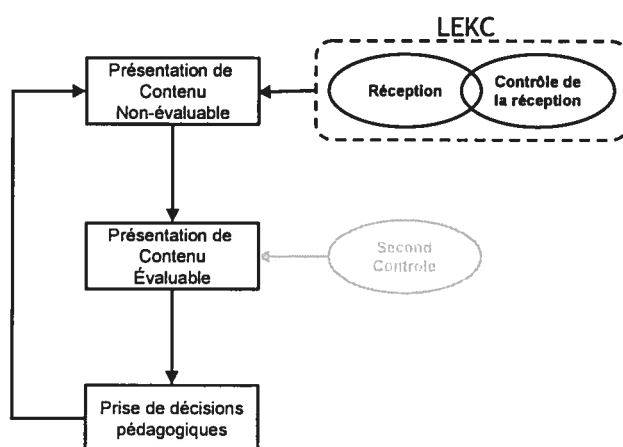


Figure 7-2 – Nouveau circuit pédagogique avec la stratégie de prévention

Avec la *stratégie de prévention*, les erreurs sont éventuellement détectées et résolues au fur et à mesure de leurs apparitions dans le contexte de la lecture et ne seront pas accumulées en attente de la phase de contrôle différée, comme c'est le cas dans les circuits pédagogiques classiques (Figure 7-1).

Ainsi, les phénomènes d'erreurs en cascade et de surcharge cognitive de l'apprenant sont évités, ou du moins réduits au maximum, ce qui augmenterait l'intérêt et la motivation de l'apprenant et influencerait positivement l'expérience d'apprentissage. D'un autre côté, les contenus évaluable comme les exercices conventionnels (e.g. question aux choix multiples, résolution de problèmes,...) peuvent toujours être utilisés comme un second niveau de contrôle.

Élargissement de la couverture des contrôles. Les contrôles dans l'approche *LEKC* sont basés sur les activités épistémiques (la reconstitution des formes épistémiques par

l'apprenant). Comme on peut virtuellement associer une forme épistémique à toute idée ou concept véhiculés dans le texte, le risque qu'une erreur de compréhension passe inaperçue est vraiment minime. La couverture des contrôles est beaucoup plus large que dans les systèmes basés sur l'approche de remédiation.

Conservation du contexte d'erreurs. Les contextes d'erreurs dans l'approche *LEKC* ne sont pas perdus. En effet, du fait de la nature du contrôle simultané (entre la réception et le contrôle de la réception), les erreurs de compréhension sont détectées et corrigées dans le contexte même où elles se sont produites. La correction de ces erreurs est grandement facilitée par deux facteurs : 1) le fait que le système possède l'information sur le contexte en cours (séquence de texte impliquée, forme épistémique en construction, base de la connaissance collective traitant ce passage,...), et 2) l'apprenant a encore toutes les données fraîches en mémoire (mémoire de travail et mémoire à court terme).

7.3 Autres avantages de l'approche *LEKC*

- Facilité de création de cours : l'approche *LEKC* est basée sur l'exploitation des contenus textuels, généralement déjà disponibles. Des livres de références dans les différents domaines de connaissance peuvent facilement être convertis en cours avancés.
- Facilité de la conception : le concepteur de cours n'est pas perdu dans sa tâche de conception. Il est guidé et orienté par le discours du document didactique. De plus l'utilisation de librairies spécialisées de patrons épistémiques accélère grandement le travail de conception.
- Enrichissement continu de la connaissance du système : le concepteur n'est pas obligé de tout modéliser avant que son système ne soit opérationnel. Une modélisation rapide pour définir les grandes lignes peut être conduite au début, puis le système peut évoluer de lui-même grâce aux interactions avec les apprenants, qui viennent enrichir sa base de connaissance collective. À l'examen de cette connaissance amassée, le concepteur peut revoir, compléter et améliorer sa conception.
- Participation et implication des apprenants : les apprenants sans le savoir participent à l'évolution du système. Leurs interactions avec le système permettent d'enrichir son comportement futur grâce à la base de connaissance collective.

- Flexibilité d'utilisation : le système peut être utilisé par les apprenants en différents modes : 1) *direct* : comme matériel supplémentaire au texte, 2) *contrôlé* : le système oriente et contrôle la lecture, et 3) *libre* : l'apprenant devient producteur de la connaissance.
- Modélisation de la connaissance vraiment indépendante du domaine : *les sciences dures*, comme la physique ou l'astronomie, tout au tant que *les sciences molles*, comme la philosophie ou la poésie, peuvent être modélisées et enseignées sur notre système.
- Modélisation de la connaissance spécialement conçue pour tenir compte des aspects cognitifs et de réflexion lors de la lecture. Elle tient compte de multiples dimensions d'un discours : propositionnelles, rhétoriques, thématiques, structurelles, didactiques,...
- Modélisation de la connaissance orientée apprenant : tout l'effort de modélisation est fait dans l'objectif de supporter la réflexion de l'apprenant et non pour supporter les inférences du système.
- Adaptabilité et extensibilité : le modèle de la connaissance proposé est adaptable aux différents domaines de connaissance. Des extensions dans le vocabulaire épistémiques (des nœuds et des liens) ainsi que dans les librairies de patrons épistémiques peuvent être opérées pour adapter le système à un domaine de connaissance en particulier.
- Transformation constructiviste de la lecture : l'approche **LEKC** est un moyen simple et efficace pour transformer la phase de lecture qui est très passive dans les systèmes actuels, en une phase active (constructiviste) où l'apprenant est très impliqué.
- Facilité d'intégration avec un SIT existant : L'idée d'utiliser un système de lecture assistée comme le système **IT-EKC** n'est pas antagoniste avec les différentes approches de construction de systèmes STIs, mais elles sont plutôt complémentaires. Le système **IT-EKC** peut fonctionner en tant que STI autonome disposant de toutes les fonctionnalités nécessaires à un STI comme nous les avons décrites au chapitre 5. Mais il peut aussi être *greffé* à un STI déjà existant au niveau des composantes d'affichage de contenu.

7.4 les limites de notre approche

- Notre approche a été étudiée et pensée pour supporter la réflexion et le raisonnement lors de la phase de la lecture. Nous pensons que cet effort doit être étendu pour supporter l'apprenant dans la phase de résolution de problèmes et dans les exercices conventionnels. Les apprenants dans cette dernière phase ont souvent besoin de revenir sur les documents (textuels) pour clarifier une idée ou chercher une information pour résoudre une étape dans un problème donné. Notre modélisation ne tient pas compte de cet aspect.
- Le discours dans un document didactique n'utilise pas que l'artefact textuel, d'autres artefacts sont utilisés : telles les images, les figures et les illustrations. Dans les documents électroniques multimédias, encore plus d'artefacts sont utilisés comme la vidéo, les animations et les simulations. Notre modèle ne tient compte que de l'artefact textuel.
Pour tenir compte des autres artefacts, une étude plus approfondie est nécessaire pour trouver les modèles de connaissance qui les régissent, semblables au modèle du discours que nous avons proposé pour l'artefact textuel.
- Sur le plan pratique beaucoup reste à faire dans le domaine de la visualisation. Les écrans de travail deviennent rapidement encombrés et difficiles à gérer. Les techniques de *zooming* et de navigation classique ne sont pas suffisantes. Elles doivent être repensées pour tenir compte des aspects conceptuels des différents modèles de connaissances (*the conceptual zooming*).

7.5 Travaux futurs

Beaucoup de voies restent à explorer suite à la réalisation de cette thèse. Parmi les plus importantes nous avons retenu les suivantes :

- Explorer de nouveaux mécanismes pour la création de méta-tags extra-documents (c'est-à-dire sans modification des documents originaux) et qui soient peu affectés par les modifications dans les documents sources.

- Etendre l'utilisation du système au support des apprenant dans les exercices et dans la résolution de problèmes. Le système doit anticiper et faciliter le retour fréquent des apprenants sur les documents didactiques dans l'exécution des exercices.
- Améliorer la visualisation par des fonctions de navigation et de zooming intelligentes. Ces fonctions doivent tenir compte de la tâche épistémique en cours et de l'état du modèle cognitif de l'apprenant pour lui présenter la connaissance sous le meilleur angle qui l'aiderait à comprendre plus facilement et à accomplir sa tâche.
- Explorer les outils d'aide à la réorganisation des formes épistémiques. Il arrive souvent que pour construire une forme épistémique donnée, l'apprenant en construit plusieurs, qui sont transformées et réorganisées avant d'arriver à la forme voulue. Cette réorganisation nécessite un grand effort de l'apprenant et elle est très consommatrice en temps. Des outils d'aide à la réorganisation sont très souhaitables.
- Rendre l'application complètement Web pour permettre une grande diffusion du système et éviter la phase d'installation de l'application.
- Utiliser des documents XHTML ou XML au lieu des documents HTML pour plus de robustesse.
- Explorer la modélisation des ressources autres que textuelles (images, vidéos, simulations, ...) toujours dans le contexte de support à la lecture.
- Etendre l'utilisation de l'approche pour d'autres domaines autres que la formation. En effet, grâce aux différentes analyses du discours que comporte notre méthode, nous pouvons facilement envisager son utilisation comme moyen de recherche, d'accès et d'analyse de l'information textuelle. Des applications dans le domaine du texte historique et dans le domaine de la connaissance corporative sont actuellement à l'étude.

BIBLIOGRAPHIE

- [Anderson, 1988] Anderson, J., *The expert module*. Dans M.C. Polson and J.J. Richardson (eds.), *Foundations of Intelligent Tutoring Systems*. Hillsdale, New Jersey, LEA. 1988.
- [Anderson, Boyle, et Yost, 1985] Anderson, J., Boyle, C., et Yost, G., *The Geometry Tutor*. The Ninth International Joint Conference on Artificial Intelligence, pp. 1-7. Los Altos, CA. 1985.
- [ARL, 2004] Applied Research Laboratory, *Training and Instructional Design*, Document en ligne à <http://www.umich.edu/~ed626/define.html>, Penn State University. 2004.
- [Barr et Feigenbaum, 1982] Barr, A., et Feigenbaum, E. A., *The Handbook of Artificial Intelligence*, 2, Kaufmann, Los Altos. 1982.
- [Ben-Ari, 1998] Ben-Ari, M., Constructivism in computer science education. Proceedings of the twenty-ninth SIGCSE technical symposium on Computer science education, pp. 257–261. 1998.
- [Biggs, 1999] Biggs, J., *Teaching for Quality Learning at University*. Buckingham: Society for Research in Higher Education and Open University Press. 1999.
- [Bloom, 1969] Bloom, B., Taxonomie des objectifs pédagogiques. Tome 1 : Domaine cognitif. Education nouvelle, Inc. 1969.
- [Bloom, 1978] Bloom, B., Taxonomie des objectifs pédagogiques. Tome 2 : Domaine affectif. Education nouvelle, Inc. 1978.
- [Brown, Burton et Bell, 1975] Brown, J.S., Burton, R.R., et Bell, A.G., *Sophie: A step towards a reactive learning environment*, International Journal of Man Machine Studies, 7, pp. 675-696. 1975.
- [Bru, 2001] Bru, M., *Approches systémiques et recherches en sciences de l'éducation*. Presses universitaires du Mirail - Les Dossiers des Sciences de L'éducation N°3. Sciences sociales. 2001.
- [Bruner, 1973] Bruner, J., *Going Beyond the Information Given*. New York: Norton. 1973.

- [Burns et Capps, 1988] Burns, H. et Capps, C., Foundations of intelligent tutoring systems dans [Polson et Richardson, 1988]
- [Burton et Brown, 1976] Burton, R., et Brown, J., *A tutoring and student modelling paradigm for gaming environments*. Dans Colman, R. and Lorton, P.Jr. (Eds.): Computer Science and Education, ACM SIGCSE Bulletin, 8(1), pp. 236-246. 1976.
- [Burton, 1982] Burton, R., et Brown, J., *An investigation of computer coaching for informal learning activities*. Dans D. Sleeman & J.S. Brown (dir.), Intelligent tutoring systems. Academic Press, New York, NY, pp. 79-98. 1982.
- [Carbonell, 1970] Carbonell, J. R., *AI in CAI: an artificial intelligence approach to computer-assisted instruction*. IEEE Transactions on Man-Machine Systems, 11, pp. 190-202. 1970.
- [Cerri, 1999] Cerri, S. A., *Shifting the focus from control to communication: the STREAMS Objects Environments model of communicating agents*. In Collaboration between Human and Artificial Societies, Coordination and Agent-Based Distributed Computing, vol. 1624, J.A. Padget, Ed. Berlin Heidelberg New York: Springer-Verlag, Lecture Notes in Artificial Intelligence, pp. 71-101. 1999.
- [Chi et al., 1989] Chi, M., Bassok, M., Lewis, M., Reimann, P., and Glaser, R. *Self-explanation: How students study and use examples in learning to solve problems*, volume 13. Cognitives Science, pp. 145-182. 1989.
- [Clancey, 1982] Clancey, W.J., *Tutoring rules for guiding a case method dialogue*. Dans Intelligent Tutoring Systems, pp. 201-225. New York, Academic Press. 1982.
- [Clancey, 1986] Clancey, W., J., *From Guidon to Neomycin and Heracles in Twenty Short Lessons*. ORN Final Report 1979-1985. AI Magazine 7(3), pp. 40-60, 187. 1986.
- [Claxton, 1998] Claxton, G., *Hare Brain, Tortoise Mind*. London: Fourth Estate. 1998.
- [Collins et Ferguson, 1993] Collins, A., et Ferguson, W., *Epistemic forms and epistemic games: Structures and strategies to guide inquiry*, Educational Psychologist, 28(1), 1993, pp. 25-42.
- [Collins et Quillian, 1969] Collins, A.M. et Quillian, M.R., Retrieval Time From Semantic Memory, *Journal of verbal learning and verbal behavior*, 8, pp. 240-248. 1969.
- [Cox et Brna, 1995] Cox, R. and Brna, P., Supporting the use of external representations in problem solving: the need for flexible learning environments. *Journal of Artificial Intelligence in Education*, 6(2), pp. 239-302. 1995.
- [Crowder, 1959] Crowder, N. A., *Automatic tutoring by means of intrinsic programming*. Automatic Teaching: The State of Art, Wiley, NY, pp. 109-116. 1959.

- [Dadong et Johnson, 1994] Dadong, W., et Johnson, P.M., *Computer supported collaborative learning using CLARE: the approach and experimental findings*. Dans les proceedings du CSCW 1994, ACM. 1994.
- [Denhière et Baudet, 1992] Denhière, G., et Baudet, S., *Lecture, compréhension de texte et science cognitive*. Presses universitaires de France (PUF), Le Psychologue. 1992.
- [Denhière, 1985] Denhière, G., *De la compréhension à lecture*, l'Orientation scolaire et professionnelle, 4, 14, pp. 305-329. 1985.
- [Dijk et Kintsch, 1983] Dijk, T. A. V. and Kintsch, W., *Strategies of discourse comprehension*. Academic Press, New York. 1983.
- [Dimitrova, Self et Brna, 1999] Dimitrova, V., Self, J., and Brna, P., The interactive maintenance of open learner models. *Proceedings of the 9th International Conference on Artificial Intelligence in Education*, pp. 405-412. 1999.
- [Du Boulay, 2000] Du Boulay, B., *Can We Learn from ITSs ? Intelligent Tutoring Systems, ITS2000*. Springer-Verlag. 2000.
- [Ertmer et Newby, 1993] Ertmer, P.A. et Newby, T.J. , *Behaviorism, Cognitivism, Constructivism: Comparing critical features from an Instructional Design perspective*. Performance Improvement Quarterly, 6(4), pp. 50-72. 1993.
- [Fisher, 1987] Fisher, W. R., *Human Communication as Narration : Toward a Philosophy of Reason, Value and Action*. University of South Carolina Press, Columbia, South Carolina. 1987.
- [Frasson et Abou-Jaoude, 1998] Frasson, C. et Abou-Jaoude, S., *Emotion Computing in Competitive Learning Environments*. Wokshop II : Pedagogical Agents, ITS-98 Conference, Fourth International Conference on Intelligent Tutoring Systems, pp. 33-39. San Antonio, Texas. 1998.
- [Gagné et al., 1992] Gagné, R.M., Briggs, L. et Wager, W., *Principles of instructional design*. Harcourt Brace Jovanovich, Orlando, FL, 4e edition. 1992.
- [Gagné, 1965] Gagné, R.M., *The Conditions of Learning*. Holt, Reinhart and Winston. 1965.
- [Gagné, 1974] Gagné, R.M., *Essentials of learning for instruction*. (2nd ed.). Hinsdale, IL: The Dryden Press. 1974.
- [Gagné, 1985] Gagné, R.M., *The Conditions of Learning*. New York: Holt Rinehart and Winston. 1985.
- [Good et Brophy, 1990] Good, T. L., Brophy, J. E., *Educational psychology: A realistic approach*. (4th ed.) White Plains, NY: Longman. 1990.
- [Grosz et Sidner, 1986] Grosz, B. et Sidner, C., *Attention, intentions, and the structure of discourse*. Computational Linguistics, 12(3), pp. 175-204. 1986.
- [Half, 1988] Half, H., *Curriculum and Instruction in Automated Tutors*. Dans [Polson et Richardson, 1988], pp. 79-108. 1988.

- [Hirschberg et Litman, 1987] Hirschberg, J., et Litman, D., *Now let's talk about now. Identifying cue phrases intonationally*. Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics (ACL-87), pp.163-171. 1987.
- [Hollan et al., 1984] Hollan, J.D., Hutchins, E.L., et Weitzman, L., *STEAMER: An interactive inspectable simulation-based training system*, AI Mag. 5 (2), pp. 15-27. 1984.
- [Jameson, 1999] Jameson, A., *User-adaptive systems: An integrative overview*. The 16th International joint conference on Artificial Intelligence (IJCAI'99). 1999.
- [Jonassen, 1995] Jonassen, D.H., *Operationalizing Mental Models: Strategies for Assessing Mental Models to Support Meaningful Learning and Design- Supportive Learning Environments*. The first international conference on Computer support for collaborative learning. Bloomington, Indiana, pp. 182-186. 1995.
- [Kaplan et Rock, 1995] Kaplan, R., et Rock, D., *New directions for intelligent tutoring*. AI Expert, pp. 31-40. 1995.
- [Khuwaja, 1994] Khuwaja Ali, R., *A Model Of Tutoring : Facilitating Knowledge Integration Using Multiple Models Of The Domain*. Thèse de doctorat en informatique, Illinois Institute of Technology. 1994.
- [Knight, 2002] Knight, P., *Complex Learning: why it matters, bat it is and how it may be stimulated and supported in higher education*. The STRATA/ETAN group on Foresight for Higher Education in Europe. 2002.
- [Laborde, 1995] Laborde, J., *Des connaissances abstraites aux réalités artificielles, le concept de micromonde Cabri*. Eyrolles, Paris. 1995.
- [Lesgold, 1988] Lesgold, A., *Toward a Theory of Curriculum for Use in Designing Instructional Systems*. Dans Mandl & Lesgold (Eds.), *Learning Issues for Intelligent Tutoring Systems*, Springer-Verlag, New York. 1988.
- [Longacre, 1983] Longacre, R., *The Grammar of Discourse*, Plenum Press, New York. 1983.
- [Luger et Stubblefield, 1998] Luger, G. and Stubblefield, W. *Artificial Intelligence, Structure And Strategies For Complex Problem Solving*. Addison-Wesley, Massachusetts, 3 edition. 1998.
- [Luger, 1994] Luger, G. F., *Cognitive Science : The Science of Intelligent Systems*. Academic Press. 1994.
- [Mann et Thompson, 1988] Mann, W., et Thompson, S., *Rhetorical structure theory: Toward a functional theory of text organization*, Text, 8(3), pp. 243-281. 1988.
- [Mann, 2003] Mann, W., *Introduction à la Théorie de la Structure Rhétorique (Rhetorical Structure Theory : RST)*. Document en ligne à <http://www.sil.org/~mannb/rst/frintro.htm> . 2003.

- [Marcu, 1997] Marcu, M., *The Rhetorical Parsing Summarization and Generation of Natural Language Texts*, Thèse de Ph.D. en informatique, Département d'informatique, Université de Toronto. 1997.
- [Mergel, 1998] Mergel, B., *Instructional Design & Learning Theory*. Document en ligne, University of Saskatchewan. 1998.
- [Merrill, 1983] Merrill, M. D., *The Component Display Theory*. Dans Reigeluth (Ed.), *Instructional Design Theories and Models: An Overview of their Current Status*. Ed. C. M., Hillsdale: Erlbaum. 1983.
- [Merrill, 1991] Merrill, M.D., *Constructivism and Instructional design*. *Educational Technology*, 31(5), pp. 45-53. 1991.
- [Merrill, 1993] Merrill, M.D., *Instructional transaction theory: Knowledge relationships among processes, entities and activities*. *Educational Technology*, 5. 1993.
- [Merrill, 2000] Merrill, M.D., *Instructional Transaction Theory (ITT): Instructional Design Based on Knowledge Objects*. Chapitre 17 dans Reigeluth (Ed.), *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory*. Mahwah, NJ: Lawrence Erlbaum Associates. 2000.
- [Merrill, 2002] Merrill, M.D., *Knowledge objects and mental models*. Chapitre 5 dans le livre "The Instructional Use of Learning Objects", Wiley, D. (ed.). Agency for Instructional Technology. 2002.
- [Mindjet, 2003] Mindjet, MindManager Business and Standard Edition. <http://www.mindjet.com/>. 2003.
- [Minsky, 1975] Minsky, M. *A framework for representing knowledge*. Dans P. Winston, Ed., *The Psychology of Computer Vision*. New York: McGraw-Hill, pp. 211-277. 1975.
- [Moore, 1995] Moore, J., *Participating in Explanatory Dialogues*. Cambridge, MA: MIT Press. (ACL-MIT Press Series). 1995.
- [Murray, 1996-a] Murray, T., *Having It All, Maybe: Design Tradeoffs in ITS Authoring Tools*. *Intelligent Tutoring Systems 1996*: pp. 93-101. 1996.
- [Murray, 1996-b] Murray, T., *Special Purpose Ontologies and the Representation of Pedagogical Knowledge*. ICLS'96, IL, USA. 1996.
- [Murray, 1998-b] Murray, T., *Expanding the Knowledge Acquisition Bottleneck for Intelligent Tutoring Systems*. *International Journal of AI and Education*, 8(3). 1998.
- [Murray, 1999] Murray, T., *Authoring Intelligent Tutoring Systems: An analysis of the state of the art*. *International Journal of Artificial Intelligence in Education*, Vol. 10, pp. 98-129. 1999.
- [Nkambou, 1996] Nkambou, R., *Modélisation des connaissances de la matière dans un système tutoriel intelligent : modèles, outils et applications*. Thèse de doctorat en informatique, Université de Montréal. 1996.

- [Novak et Gowin, 1984] Novak, J.D. , et Gowin, D.B., *Learning How To Learn*. New York, Cambridge University Press. 1984.
- [Nwana, 1990] Nwana, H., *Intelligent Tutoring Systems: An Overview*. Artificial Intelligence Review, 4, pp. 251-277. 1990.
- [Ohlsson , 1994] Ohlsson, S., *Constraint-Based Student Modelling*. Student Modelling: The Key to Individualized Knowledge-Based Instruction, pp. 167-189, Springer-Verlag. 1994.
- [Ohlsson, 1992] Ohlsson, S., *Constraint-Based Student Modelling*. Artificial Intelligence in Education, 3(4), pp. 429–447. 1992.
- [Ohlsson, 1996] Ohlsson, S., *Learning to do and learning to understand*. Dans Reimann, P. and Spada, H. (Eds.), *Learning in humans and machines*, Pergamon, oxford edition. 1996.
- [Patel et Kinshuk, 1997] Patel, A. et Kinshuk, *Intelligent Tutoring Tools in a Computer Integrated Learning Environment for introductory numeric disciplines*. Innovations in Education and Training International Journal, 34(3), pp. 200–207. 1997.
- [Perkins et Unger, 1999] Perkins, D.N., et Unger, C., “Teaching and Learning for Understanding”, in Reigeluth, C.M., *Instructional-Design Theories and Models New Paradigm of Instructional Theory*, Volume II, LEA, NJ, pp. 91-114. 1999.
- [Perkins, 1991] Perkins, D. N., *Technology meets constructivism: Do they make a marriage?* Educational Technology, 31(5), pp. 18-23. 1991.
- [Polanyi, 1988] Polanyi, P., *A formal model of the structure of discourse*, Journal of Pragmatics, 12, pp. 601-638. 1988.
- [Polson et Richardson, 1988] Polson, M. and Richardson, J., *Foundations of Intelligent Tutoring Systems*. LEA, Hillsdale, New Jersey. 1988.
- [Polson, 1993] Polson, M.C., Cognitive Theory as a Basis for Instructional Design. Chapitre 1 dans [Spector, Polson et Muraida, 1993], pp. 5-22. 1993.
- [Prince, 1992] Prince, E.F., The ZPG letter: Subjects, definitions, and information-status. In W.C. Mann & S. Thompson (eds), *Discourse Description: Diverse Linguistic Analysis of a Fund-Raising Text*, pp. 295-325. John Benjamins Publishing Co. 1992.
- [Quéré, 1991] Quéré, M., *Système Experts et Enseignement Assisté par Ordinateur*. A.E.M. OPHRYS. 1991.
- [Ragnemalm, 1995] Ragnemalm, E. L. , *Student Diagnosis in Practice ; Bridging a Gap*. International Journal of User Modeling and User-Adapted Interact, 5(2), pp. 93–116. 1995.
- [Reigeluth et al., 1978] Reigeluth, C.M., Merrill, M.D. et Bunderson, C.V. , *The structure of subject matter content and its instructional design implications*. Instructional Science, 7(2), pp. 107-126. 1978.

- [Reigeluth et Moore, 1999] Reigeluth, C.M., et Moore, J., *Cognitive Education and The Cognitive Domain*. Dans Reigeluth, C.M., *Instruction-Design Theories and Models, A new Paradigm of Instructional Theory*, Volume II, pp. 51-68. 1999.
- [Reimann, 1991] Reimann, P. *Detecting functional relations in a computerized discovery environment*. *Learning and Instruction*, 1, pp. 45–65. 1991.
- [Reiser, Anderson et Farrell, 1985] Reiser, B.J., Anderson, J.R., et Farrell, R.B., *Dynamic student modeling in an intelligent tutor for LISP programming*. *International Joint Conference on Artificial Intelligent-85*, vol. 1, pp. 8-14. Los Altos, CA. 1985.
- [Resnick, 1981] Resnick, L. B., *Instructional psychology*. *Annual Review of Psychology*, 32, pp. 659-704. 1981.
- [Rodenburg, 2001] Rodenburg, D., *Learning System Design: More than Atomic Science*. *Learning Technology, IEEE Computer Society (LTTF)*, 3(1). 2001.
- [Rouane, 1999] Rouane, K., *Étude et réalisation d'un environnement de développement de curriculum*. Mémoire de maîtrise, Université de Montréal, Québec. 1999.
- [Rouane, Frasson et Kaltenbach, 2002-a] Rouane, K., Frasson, C., et Kaltenbach, M., *LKC: Learning by Knowledge Construction*. *Intelligent Tutoring Systems 2002*, Biarritz, France, pp. 188-198. 2002.
- [Rouane, Frasson et Kaltenbach, 2002-b] Rouane, K., Frasson, C., et Kaltenbach, M., *Advanced Annotation and External Representation in LKC System*. *Technology of Information and Communication in Education for Engineering and Industry, TICE'2002*, Lyon, France. 2002.
- [Rouane, Frasson et Kaltenbach, 2003-a] Rouane, K., Frasson, C., et Kaltenbach, M., *Reading for Understanding: A Framework for Advanced Reading Support*. *The 3rd IEEE International Conference on Advanced Learning Technologies*, Athens, Greece, pp. 9-11. 2003.
- [Rouane, Frasson et Kaltenbach, 2003-b] Rouane, K., Frasson, C., et Kaltenbach, M., *A Framework for an Advanced Reading Support in the Digital Library Age*. Dans *The 2003 IEEE/WIC International Conference on Web Intelligence (WI 2003)*. October 13-16, 2003, Halifax, Canada. 2003.
- [Rumelhart et Norman, 1978] Rumelhart, D. E., et Norman, D. A. *Accretion, Tuning, and restructuring: Three modes of learning*. Dans J. W. Cotton, et R. L. Klatzky (Eds.), *Semantic factors in cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates. 1978.
- [Saettler, 1990] Saettler, P., *The Evolution of American Educational Technology*, Englewood, CO:Libraries Unlimited, Inc. 1990.
- [Sandberg et Andriessen, 1997] Sandberg, J., et Andriessen, J., *Where is AI and how about education ?* *The 8th AIED'97*. Kobe, Japan. 1997.

- [Sandberg et Andriessen, 1999] Sandberg, J., et Andriessen, J., *Where is education heading and how about AI ?* International Journal of Artificial Intelligence in Education, 10(2), pp. 130-150. 1999.
- [Seidel et Park, 1994] Seidel, R., et Park, O., An historical perspective and a model for evaluation of intelligent tutoring systems. *Journal of Educational Computing Research*, 10(2), pp. 103-128. 1994.
- [Self, 1974] Self, J. A., *Student models in computer-aided instruction*. International Journal of Man-machine Studies, 6, pp. 261-276. 1974.
- [Self, 1988] Self, J.A., *Knowledge, Belief and User Modelling*. AIMSAS'1988. pps 3-9. 1988.
- [Self, 1990] Self, J., *Bypassing the Intractable Problem of Student Modelling*. Dans Frasson, C. and Gauthier, G. (eds.), *Intelligent Tutoring Systems: at the Crossroads of Artificial Intelligence and Education* Norwood, NJ, pp. 107-123. 1990.
- [Self, 1990-a] Self, J., *Theoretical foundations of intelligent tutoring systems*. Journal of Artificial Intelligence in Education, 1(4), pp. 3-14. 1990.
- [Self, 1999] Self, J. A., *The defining characteristics of intelligent tutoring systems research: ITSs care, precisely*. International Journal of Artificial Intelligence in Education, 10, pp. 350-364. 1999.
- [Shortliffe, 1976] Shortliffe, E., *Computer-based medical consultations: MYCIN*. New York Academic Elsevier. 1976.
- [Skinner, 1958] Skinner, B. F. , *Teaching Machines*. Science, 128, pp. 969-977. 1958.
- [Smith et al., 1974] Smith, E., Shoben, E., et Rips, L., *Structure and process in semantic memory: A feature model for semantic decisions*. Psychological Review, 81, pp. 214-241. 1974.
- [Symes et McIntyre, 2000] Symes, C. et McIntyre, J. , *Working Knowledge: the new vocationalism and higher education*. Buckingham: Society for Research in Higher Education and Open University Press. 2000.
- [Tadié, 1998] Tadié, G., *Un système multi-agent pour l'enseignement et la simulation de tâches coopératives*. Thèse de doctorat en informatique, Université of Montréal, Montréal, Canada. 1998.
- [Tarnas, 1991] Tarnas, R., *The passion of the western mind*. New York: Harmony Books. Part VI: The transformation of the modern era. 1991.
- [Toulmin, 1958] Toulmin, S., *The uses of argument*. Cambridge, MA: Cambridge University Press. 1958.
- [Twitchell, 1990] Twitchell, D., Série de coresponsables entre Gagné R.M. et Merrill, D.. Educational Technology. 1990.
- [Uhr, 1969] Uhr L., *Teaching machine programs that generate problems as a function of interaction with students*. The 24th National Conference of Education, Virginia, April 2-4, 1969.

- [VanLehn, 1988] VanLehn, K., *Student modeling*, dans Polson & Richardson, pp. 55-77. 1988.
- [Vasandani et Govindaraj, 1995] Vasandani, V., et Govindaraj, T., *Knowledge organization in intelligent tutoring systems for diagnostic problem solving in complex dynamic domains*. IEEE Transactions on System, Man, and Cybernetics, 25(7), pp. 1076-1096. 1995.
- [Vosniadou et Brewer, 1992] Vosniadou, S., et Brewer, W.F.,. *Mental models of the earth: A study of conceptual change in childhood*. Cognitive Psychology, 24, pp. 535-85. 1992.
- [Vosniadou, 1996] Vosniadou, S., *A cognitive psychological approach to learning*. Dans Reimann, P. and Spada, H. (Eds.), *Learning in humans and machines*, Pergamon, Oxford edition. 1996.
- [Vosniadou, 2002] Vosniadou, S., *Mental Models in Conceptual Development*. Dans L. Magnani et N. Nersessian, *Model-Based Reasoning: Science, Technology, Values*, New York: Kluwer Academic Press. 2002.
- [Weller, 2000] Weller, M. , *The Use of Narrative to Provide a Cohesive Structure for a Web Based Computing Course*. Journal of Interactive Media in Education. 2000.
- [Weller, 2000] Weller, M., *The Use of Narrative to Provide a Cohesive Structure for a Web Based Computing Course*. Journal of Interactive Media in Education. 2000.
- [Wenger, 1987] Wenger, E., *Artificial Intelligence and Tutoring Systems*. Morgan Kaufman. 1987.
- [White et Frederiksen, 1985] White, B. Y. , et Frederiksen, J., *QUEST:Qualitative understanding of Electrical troubleshooting*, ACM SIGART Newsletter, 93, pp. 34-37. 1985.
- [Wilson et al., 1995] Wilson, B., Teslow, J., et Osman-Jouchoux, R., *The Impact of Constructivism (and Postmodernism) on ID Fundamentals*. In B. B. Seels (Ed.), *Instructional Design Fundamentals: A Review and Reconsideration*, pp. 137-157. Englewood Cliffs NJ: Educational Technology Publications. 1995.
- [Wilson et Cole, 1996] Wilson, B. G., et Cole, P., *Cognitive teaching models*, dans Handbook of research in instructional technology, David H. Jonassen, Editor, New York: Scholastic Press. 1996.
- [Winograd et Flores, 1987] Winograd, T., et Flores, F., *Understanding computers and cognition: A new foundation for design*, Addison-Wesley Publishing Company. 1987.
- [Zhou et Events, 1999] Zhou, Y. et Evens, M. W., *A Practical Student Model in an Intelligent Tutoring System*. Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence, Chicago, IL, pp. 13-18. 1999.

ANNEXE A

MODELISATION THEORIQUE DE *LEKC*

Soit un document \mathcal{T} représentant le texte source sous la forme d'une chaîne de caractères $[c_1 c_2 \dots c_n]$ où $c_{i=1..n}$ est un caractère.

Augmenter le document \mathcal{T} c'est lui adjoindre le modèle de connaissance $LEKC(\mathcal{T})$. Ce qui revient à définir ses modèles du discours, épistémique et didactique, respectivement notés $LEKC-MD(\mathcal{T})$, $LEKC-ME(\mathcal{T})$ et $LEKC-MI(\mathcal{T})$.

$$LEKC(\mathcal{T}) = [LEKC-MD(\mathcal{T}), LEKC-ME(\mathcal{T}), LEKC-MI(\mathcal{T})]$$

A.1 Quelques définitions

- $t[c, \dots c]$: est une sous-chaîne extraite de \mathcal{T} .
- $\mathcal{P}(t)$: est l'ensemble des sous-chaînes $t[c, \dots c]$ extraites de \mathcal{T} .

- \cap : Intersection de chaînes de caractères. Pour des chaînes $m1$ et $m2$ de \mathcal{T} en supposant que $m1$ apparaît avant $m2$ dans \mathcal{T} :

$$m1[m_{a1} .. m_{ak}] \cap m2[m_{b1} .. m_{bl}] = \begin{cases} t[m_{b1} .. m_{ak}] & \text{si } ak > b1 \\ \emptyset & \text{sinon} \end{cases}$$

A.2 Le modèle du discours $LEKC-MD(\mathcal{T})$

- $m\mathcal{A}[c, \dots, c]$: est une microproposition *atomique* de \mathcal{T} :

$$\begin{cases} m\mathcal{A} \in \mathcal{P}(t) \text{ et} \\ \forall m\mathcal{A}_i \in \mathcal{P}(t), \text{ si } m\mathcal{A}_i \neq m\mathcal{A} \text{ alors } m\mathcal{A}_i \cap m\mathcal{A} = \emptyset \end{cases}$$

- $\mathcal{P}(m\mathcal{A})$ est l'ensemble des micropropositions *atomiques* extraites de \mathcal{T} .
- $m\mathcal{C}[m_1, \dots, m_k]$ est une microproposition *composée* :

$$\forall m_{i=1..k}, m_i \in \mathcal{P}(m\mathcal{A}) \cup \mathcal{P}(m\mathcal{C})$$

- $\mathcal{M}a(\text{Opération}, \text{Proposition}, [m_1, \dots, m_k])$ est une macroproposition définie par :

$$\begin{cases} \text{Opération} \in \{\text{Généralisation}, \text{Construction}\} \\ \text{Proposition} \text{ est une chaîne de caractère} \\ m_1, \dots, m_k \in \mathcal{P}(m\mathcal{A}) \cup \mathcal{P}(m\mathcal{C}) \cup \mathcal{P}(\mathcal{M}a) \end{cases}$$

- $LEKC-MD(\mathcal{T})$ le modèle du discours est défini par :

$$LEKC-MD(\mathcal{T}) = [\mathcal{P}(m\mathcal{A}), \mathcal{P}(m\mathcal{C}), \mathcal{P}(\mathcal{M}a)]$$

A.3 Le modèle épistémique $LEKC-ME(\mathcal{T})$

- $\mathcal{FÉ}$: une forme épistémique est un graphe $\langle \mathcal{N}, \mathcal{A} \rangle$ défini par :

$$\mathcal{N} \subset \mathcal{P}(m\mathcal{A}) \cup \mathcal{P}(mC) \cup \mathcal{P}(Ma) \cup \mathcal{P}(t\mathcal{H}) \cup \mathcal{P}(s\mathcal{T}) \cup \mathcal{P}(\mathcal{FÉ})^*$$

$$\mathcal{A} = \{ \langle r, [n_1, n_2, \dots, n_k] \rangle \mid r \in \mathcal{P}(r\mathcal{R}) \cup \mathcal{P}(r\mathcal{T}) \cup \mathcal{P}(rS) \text{ et } [n_1, n_2, \dots, n_k] \in \mathcal{N} \times \mathcal{N} \times \dots \times \mathcal{N} \}$$

- $t\mathcal{H}$: un thème est défini par :

$$t\mathcal{H} = \langle h, \text{ÉlémentProposition} \rangle \text{ ou } h \text{ est un type thématique}$$

$$\text{ÉlémentProposition} \in \mathcal{P}(\mathcal{FÉ})$$

Notons que $\mathcal{P}(\mathcal{FÉ})$ inclut $\mathcal{P}(m\mathcal{A})$, $\mathcal{P}(mC)$, $\mathcal{P}(Ma)$, $\mathcal{P}(t\mathcal{H})$ et $\mathcal{P}(s\mathcal{T})$

$$h \in \{ \textit{Affirmation, Conjecture, Hypothèse, Fait, Axiome, Concept, Méthode, Preuve, Problème, Question, Théorie} \}$$

- $s\mathcal{T}$: représente une construction structurelle :

$$s\mathcal{T} = \langle s, [n_1, n_2, \dots, n_j] \rangle \mid n_{i=1..j} \in \mathcal{N} \text{ et}$$

$$s \in \{ \textit{Liste, Séquence, Comparaison, Conjonction, Disjonction} \}$$

- $r\mathcal{R}$ est l'ensemble des relations rhétoriques :

$$r\mathcal{R} = \{ \textit{Antithèse, Concession, Arrière-plan, Préparation, Facilitation, Démonstration, Justification, Reformulation, Résumé, But, Cause-délibérée, Cause-non-délibérée, Motivation, Circonstance, Évaluation, Interprétation, Élaboration, Moyen, Résultat-délibéré, Solution, Condition, Anti-condition, Inconditionnel, À-moins-que, Résultat-non-délibéré} \}$$

* Les nœuds dans \mathcal{N} peuvent être eux mêmes des forme épistémique de $\mathcal{P}(\mathcal{FÉ})$ qui est l'ensemble des formes épistémiques.

- $r\mathcal{T}$ est l'ensemble des relations thématiques :

$$r\mathcal{T} = \{ \textit{Affaiblit}, \textit{Renforce}, \textit{Alternative-à}, \textit{Supporte}, \textit{Contredit}, \textit{Répond-à}, \textit{Réfère-à}, \textit{Suggère} \}$$

- $r\mathcal{S}$ est l'ensemble des relations structurelles :

$$r\mathcal{S} = \{ \textit{Est-un(e)}, \textit{Est-membre-de}, \textit{Constitué-de}, \textit{Composé-de}, \textit{Contraste} \}$$

- $LEKC-ME(\mathcal{T})$ le modèle épistémique est défini par :

$$LEKC-ME = \mathcal{P}(\mathcal{F}\mathcal{E})$$

A.4 Le modèle didactique $LEKC-MI(\mathcal{T})$

- $s\mathcal{G}$: est un sujet génératif.

- $rs\mathcal{G}$: est l'ensemble de relations d'organisation des sujets génératifs.

$$rs\mathcal{G} = \{ \textit{Pré-requis}, \textit{Réfère-à}, \textit{Composé-de} \}$$

- $o\mathcal{C}$: est un objectif de compréhension.

- $ro\mathcal{C}$: est l'ensemble de relations d'organisation des objectifs de compréhensions.

$$ro\mathcal{C} = \{ \textit{Pré-requis}, \textit{Alternative-à}, \textit{Sujet-Objectif}, \textit{Composé-de} \}$$

- $a\mathcal{E}$: est une activité épistémique.

- $ra\mathcal{E}$: est l'ensemble de relations d'organisation des activités épistémiques.

$$ra\mathcal{E} = \{ \textit{Pré-requis}, \textit{Alternative-à}, \textit{Objectif-Activité}, \textit{Cible} \}$$

– Le modèle didactique *LEKC-MI*(\mathcal{T}) est défini comme un graphe $\langle \mathcal{V}, \mathcal{L} \rangle$ définie par :

$$LEKC-MI(\mathcal{T}) = \langle \mathcal{V}, \mathcal{L} \rangle$$

$$\mathcal{V} \subset \mathcal{P}(s\mathcal{G}) \cup \mathcal{P}(o\mathcal{C}) \cup \mathcal{P}(a\mathcal{E})$$

$$\mathcal{L} = \{ \langle r, [n1, n2] \rangle \mid r \in rs\mathcal{G} \cup ro\mathcal{C} \cup ra\mathcal{E} \text{ et } [n1, n2] \in \mathcal{V} \times \mathcal{V} \}$$

